UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Temporal interactions of microbiota in longitudinal nasopharyngeal samples and association with lower respiratory tract infection

Brian Thabane Rambau (RMBBRI001)

Dissertation submitted to
the School of Public Health and Family Medicine,
University of Cape Town,
in partial fulfilment of the requirements for the
Master of Public Health (Epidemiology & Biostatistics) degree

Supervisor: Dr. Maia Lesosky

September 2019

DECLARATION

I, Brian Thabane Rambau (RMBBRI001), hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate Date: 20 September 2019

DEDICATION

To Senior Prophet T.B Joshua, whose vision of Emmanuel TV has been a saving grace, and to my biological parents Julia and Elias for their encouragement throughout the hard times.

All glory to God who gave the grace for completion of this thesis.

ACKNOWLEDGEMENTS

I would like to acknowledge my supervisor, Dr. Maia Lesosky, for your support and the rarest opportunity I was afforded.

I thank also Prof M. Nicol for the opportunity to work with the DCHS data and for the conception of the research idea.

I was supported during my studies by the Medical Research Council of South Africa in terms of the National Health Scholars Programme from funds provided for this purpose by the National Department of Health/Public Health Enhancement Fund and by the South African Department of Science and Technology/National Research Foundation (DST-NRF), Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, Stellenbosch, South Africa.

DISSERTATION ABSTRACT

During aetiology of respiratory illnesses, it is widely accepted that infection is preceded by nasopharyngeal (NP) colonisation with bacteria and that NP flora develop early in childhood (during the first year of life). The presence of multiple NP bacteria results in competitive and synergistic associations, however temporal organism interactions have rarely been explored due to limited availability of longitudinal data sets, and the complex statistical methods needed. This study aimed to identify, describe and quantify the temporal interactions existing between selected key bacteria colonizing the nasopharynx in young children (up to 1 year old), and to further compare these patterns in children who go on to develop pneumonia compared to those who do not.

The significance of the study, as well as the objectives of the study, methods and data analysis plan are outlined in the study protocol (Part A). A summary of what is currently known about NP bacterial species interactions is presented as part of the literature review (Part B). The primary aim of the literature review was to describe the prevalence of NP carriage of four NP colonizing bacteria of interest: S. *pneumoniae*, S. *aureus*, H. *influenzae* and M. *catarrhalis* in children, as well as identify any risk factors or confounding associations. The literature review furthermore aimed to identify previously described NP bacterial species interaction patterns, as well as providing a summary of statistical approaches previously employed in the studying bacterial interactions. A manuscript presenting the subsequent analysis of these data is included as Part C.

This study was a secondary data analysis of 760 infants enrolled in a birth cohort with NP swabs collected every two weeks for the first year of life and additionally at episodes of lower respiratory tract infections (LRTI). Kaplan-Meier estimates were used to visualize time to first carriage. Generalised estimating equations with a logit link and adjusted for repeated measures were used to estimate the time varying association of NP bacteria carriage with development of pneumonia, while enabling adjustment by key confounders. Markov multi state models (MSMs) were used to describe NP bacterial acquisition with age and estimation of clearance probabilities, new acquisition or persistent acquisition.

There were 760 individuals included in the analysis, with a total of 16,346 NP samples available and a median 364 person-days (IQR 346 – 365 person-days). S. *pneumoniae* was predominant, found in >55% of all samples and demonstrating carriage in >95% of individuals at least once by 12 months of age. S. *aureus* was both less common (25% of

samples and 88% of individuals) but also had a strikingly different pattern of first acquisition compared to the other three organisms, demonstrating a rapid increase in carriage prevalence until approximately four weeks and subsequently decreasing.

S. *pneumoniae* had the highest co-carriage prevalence overall with H. *influenzae* and M. *catarrhalis* (both 25%) but this varied by age category. In contrast, co-carriage with S. *aureus* was less prevalent with either S. *pneumoniae* (12%), H. *influenzae* (5%) or M. *catarrhalis* (6%). Co-carriage frequencies differed considerably by age category, at least partially reflecting the relative prevalence of carriage by age. Carriage and co-carriage rates were similar among those children that experienced LRTI compared to those that did not. Seasonal carriage varied, but to a small extent compared to variance by age.

Models adjusting for sex, site, season of birth and age found temporally sustained positive associations between the co-carriages of S. *pneumoniae* with H. *influenzae*, and M. *catarrhalis*, but no association with S. *aureus*.

Clear differences occur in the co-carriage patterns of S. *pneumoniae* with other organisms. The probability of acquisition of S. *pneumoniae* is modified by earlier carriage of H. *influenzae* or M. *catarrhalis*. Positive H. *influenzae* carriage increases the probability of acquisition of S. *pneumoniae* with transition probabilities from 0.15 (95% CI 0.14-0.17) to 0.36 (95% CI 0.17, 0.54) after 28 days of age, compared to the same period probability of acquisition of S. *pneumoniae* alone at 0.015 (95% CI 0.043-0.076) to 0.088 (95% CI 0.075-0.10). There is no difference in the clearance of S. *pneumoniae* related to H. *influenzae* carriage, but clearance of H. *influenzae* before 6 months of age is far less likely if coming from a state of co-carriage (probability between 0.04 - 0.07) compared to sole carriage (probability 0.23 - 0.12). The only evidence of differences in clearance probability in the models investigating S. *pneumoniae* and M. *catarrhalis* are in the probability of M. *catarrhalis* clearance before 28d which is 0.24 (95% CI 0.15 - 0.38) if carried alone and only 0.058 (55%CI 0.01 - 0.30) if carried with S. *pneumoniae*, though these confidence intervals overlap.

Through this modelling we found positive sustained interactions between S. *pneumoniae* and both H. *influenzae*, and M. *catarrhalis*, where models indicated that preceding carriage or colonisation with either H. *influenzae*, and M. *catarrhalis* may increase the risk of colonisation with S. *pneumonia*. Timing of carriage and overall prevalence of carriage are in line with other findings in similar populations with overall high exposure to S. *pneumoniae*,

H. *influenzae*, M. *catarrhalis* during the first year of life and rapid and early exposure to S. aureus. Carriage, co-carriage and transition frequency did not vary appreciably when comparing children who experienced LRTI in the first year of life compared to those who did not, suggesting that overall exposures are similar, but that further modelling is required to understand the specific timing of associations in relations to LRTI.

LIST OF ABBREVIATIONS

AOM	Acute otitis media
CRFs	Case report forms
DCHS	Drakenstein child health study
GEE	Generalised estimating equations
HI	Haemophilus *influenzae*
Hib	Haemophilus *influenzae* type-b
HIV	Human immunodeficiency virus
LMICs	Lower and middle income countries
LRT	Lower respiratory tract
LRTI	Lower respiratory tract infection
MC	Moraxella *catarrhalis*
MSMs	Multi state models
NP	Nasopharyngeal
OM	Otitis media
PCR	Polymerase chain reaction
PI	Principal Investigator
PRSP	Penicillin-resistant Streptococcus *pneumoniae*
PTB	Pulmonary Tuberculosis
qPCR	Quantitative polymerase chain reaction
RSV	Respiratory syncytial virus
SA	Staphylococcus aureus
SD	Standard deviation
SES	Socio-economic status
URI	Upper respiratory infection
URT	Upper respiratory tract
URTI	Upper respiratory tract infection
SP	Streptococcus *pneumoniae*
USA	United States of America
WHO	World Health Organization

WORD COUNT AND REFERENCES

Abstract: 897
Protocol: 2267
Protocol references: 571
Literature review: 4739
Literature review references: 1348
Manuscript: 4336
Manuscript references: 599
Whole thesis: 23 746

TABLE OF CONTENTS

TABLE OF FIGURES

Part B: Literature review

Part C: Manuscript

TABLE OF TABLES

Part B: Literature review

Part C: Manuscript

Part A: Protocol

Protocol Title: Identifying and quantifying organism interactions in longitudinal child health studies.

Introduction

In the aetiology of bacterial pneumonia and otitis media (OM), it is generally accepted that infection is preceded by nasopharyngeal (NP) colonization (Dunne et al. 2013). Several studies support that NP flora develop and establish early in childhood or in first year of life. Of the organisms that colonize the nasopharynx (i.e. those that develop and establish), the commonly studied are S. *pneumoniae*, H. *influenzae*, M. *catarrhalis* and S. *aureus*, all of which are associated with respiratory disease in children. However, colonization of the nasopharynx of young children is often asymptomatic (Murphy et al. 2009), and organism colonisation is known to co-occur. Colonisation timing or pattern can be impacted by risk factors, antimicrobial treatment and/or vaccination (Xu et al. 2012).

A body of literature exist on organism interactions, however temporal organism interactions have not been explored, especially regarding whether observed colonization patterns are resultant of the host's immunity or synergistic and antagonistic organism interactions. An understanding of the time dependent interactions in young children could provide insight into interpretation of estimated nasopharyngeal (NP) colonization thresholds and diagnosis of respiratory tract infections.

Research aims and objectives

The primary aim of this study is to identify, describe and quantify temporal interactions existing between selected pathogens occurring in the nasopharynx in young children (up to 1 year of life), and to compare these patterns in children who have an occurrence of lower respiratory tract infection (LRTI) in the first year of life compared to those that do not.

Specific objectives:

1. Describe the frequency, seasonality, and age related patterns of S *aureus*, S *pneumoniae*, H *influenzae* and M *catarrhalis* carriage from NP samples in infants less than 1 year.
2. Estimate the direction and magnitude of co-occurrence and temporal interactions between S *pneumoniae* and each of S *aureus*, H *influenzae* and M *catarrhalis*.
3. Compare the patterns of organism colonisation between infants who have experienced an episode of LRTI in the first year of life versus those who have not.

Background

Lower respiratory tract infections (LRTI) persist as the leading cause of morbidity and mortality in young children with an estimated incidence of about 0.22 episodes per child-year in low- and middle-income countries (LMICs) in 2010 (Rudan et al. 2013). Although LRTI does not follow a one-cause one-disease postulate, there are known more common disease causing organisms. Among pneumonia cases in a South African setting, respiratory syncytial virus (RSV) was the most common pathogen followed by influenza, at 29% and 17%, respectively. Moreover, the estimated incidence of community acquired pneumonia was 0.14 (705 554/5 041 132) for all acute lower respiratory infections (LRTI) and 0.4 (203 482/5 041 132) for RSV related pneumonia in the same setting (Rudan et al. 2013). Among young children globally, at least 50% are expected to be colonized with S *pneumoniae*, H *influenzae* or M *catarrhalis* at a given point in time and approximately 35% with S *aureus* (Pettigrew et al. 2008). All four organisms are known to cause disease, primarily respiratory tract infections, in early childhood, and so represent important species of study. Though organism carriage in the nasopharynx (NP) is not necessarily indicative of colonisation, nor of progression to disease, carriage is a necessary precursor (Anthony et al. 2012).

Studying carriage risk factors may offer an opportunity to intervene on respiratory disease by enabling one to alter or modulate exposure to predisposing factors, especially in pathogenesis of respiratory infections. Risk factors for NP colonisation by one organism may be similar or different to NP colonisation by other common NP colonisers. For instance, Pettigrew et al. (2008) showed that age, race and day care attendance were associated with colonisation by H *influenzae* but not with S *pneumoniae* or M *catarrhalis*. In addition to environmental and clinical risk factors, another factor that may alter progression to disease is co-occurrence or competing occurrence of organisms. Organism interactions have been studied for years including exploration of host, agent and environmental factors, i.e. whether or not these predispose one to nasopharyngeal (NP) colonisation and subsequent disease. Lewnard et al. (2016) was interested in whether observed organism interactions were a true reflection of interspecies interactions or merely just confounded results. Using longitudinal data of children aged between 2 – 30 months, the authors arrived at the conclusion that observed organism interactions are more likely to be due to interspecies interaction rather than simple confounding either by age differences or seasonal variations.

Previous work (Heinsbroek et al. 2015) has reported distributions in carriage of S *pneumoniae* among infants by maternal HIV exposure status (no difference), age (64.7% at 14 weeks of age compared to 27.1% at 52 weeks of age) and season (August peak), using longitudinal household data. Similarly, both S. *pneumoniae* and H. *influenzae* carriage have been report to peak during winter seasons (Lewnard et al. 2016). Further work by Lewnard et al. (2016) investigating possible confounding in analysis of organism interactions, found no confounding by age or season, but potential confounding by differences in socio-economic status (SES). In their data, investigators found that among a low SES population both S *pneumoniae* and H *influenzae* acquisition was earlier and with higher overall prevalence than compared to a high SES population at two months of age, noting that the groups came from different cultural backgrounds. Co-colonisation with S. *pneumoniae* and H. *influenzae* resulted in a 25% greater decline in S. *aureus* carriage in the children from the low SES households compared to the high SES households (Lewnard et al. 2016).

A number of studies investigating NP organism interactions have been done, but often describe contradictory findings (Murphy et al. 2009). Authors have described synergism between *S. pneumoniae* + H. *influenzae* and antagonism between both S *pneumoniae* + S *aureus* and H. *influenzae* + S *aureus* (Lewnard et al. 2016). Pettigrew et al (2008), however, describe an antagonistic association between H *influenzae* and S *pneumoniae*. Although there are slight age differences in the children between the studies, that alone does not appear to sufficiently explain the different findings. Further, the results suggest that additional organisms, for example, considering co-colonisation by M *catarrhalis*, substantially increased the complexity of evaluation. In Pettigrew et al (2008), it was found that when M *catarrhalis* co-colonised with H *influenzae* the antagonism between H *influenzae* + S *pneumoniae* reversed to synergism. Inverse associations were also observed between both S *pneumoniae* and H *influenzae* individually with S *aureus* (Pettigrew et al 2008). The degree of complexity in understanding even interactions between a few organisms calls for careful longitudinal studies and considered analysis.

Most of the literature estimating organism interactions are cross sectional study designs, with only a handful being longitudinal or repeated measures designs (Jacoby et al. 2007; Pettigrew et al. 2008; Rupa et al. 2012). While longitudinal designs have obvious strengths in assessing

temporal organism interactions, accounting for the variance from repeated individual participant samplings (i.e. serial observations or correlation) is important. Previously applied methods have varied, but primarily have been focused on generalized estimating equation (GEE), a 'population-averaged' model. For example, Pettigrew et al., (2008) followed infants 6 – 36 months old for a period of 1 year and used a GEE framework under a repeated measures logistic regression with an autoregressive correlation structure to estimate population odds ratios. Similarly, Rupa et al. (2012) modelled longitudinal data accounting for serial observations through an exchangeable correlation structure, again using a GEE framework. Unlike the GEE marginal models, Jacoby et al. (2007) applied a conditional multivariable mixed effects models to estimate pairwise interactions. Apart from GEE, exploratory analysis computing carriage frequencies are part of the early literature on NP colonisation and organism interactions. With cross-sectional data, interactions are typically estimated by unadjusted odds ratios from chi-squared test or Fisher's exact test. The timing of carriage events has also being explored by survival analysis methods.

Methodology

Study Design

This data for this secondary analysis is taken from the Drakenstein Child Health Study (DCHS) (Zar et al. 2014). DCHS is a population-based birth cohort study investigating early-life determinants of child health. Mother-infant pairs were enrolled at the time of infant birth and followed longitudinally for five years. The DCHS cohort includes >1000 mother-infant pairs, and follow-up is ongoing. The enrolled participants attended routine study visits at scheduled time points as well as unscheduled visits during sick episodes. A subset of the cohort elected to be enrolled into an intensively sampled group, who underwent 2-weekly visits during the first year of life. The samples for this analysis are from the intensively sampled sub-cohort.

Characteristics of the study population

The DCHS parent study recruited pregnant mothers in their second trimester a from peri-urban area in Cape Town, characterized by low socio-economic status (SES), informal housing, crowding and a high unemployment rate (Zar et al. 2015). Neonates and their mothers are followed for at least 5 years with some enrolled into an intensively sampled early

life cohort. Children received PCV13 and Hib immunizations from the primary health care, which are part of the South African routine immunization programme with the majority (>99%) of children receiving the full course of vaccination. HIV prevalence among mothers is 25% with only 2 HIV infected children (Zar et al. 2013).

Recruitment and enrolment

Research procedures and data collection methods

This study investigates selected measures from the subset of >800 infants participating in an intensively measured cohort during the first year of life. Study visits were conducted by trained study staff and data entered into REDCap database using standardised case report forms (CRFs). Primary outcome measures are from nasopharyngeal (NP) swabs that were taken at both routine study visits (every 2 weeks) and at episodes of lower tract respiratory infection (LRTI) identified by active surveillance. Collected NP swabs (Copan flocked swab, FLOQSwabsTM, COPAN Diagnostics, Murrieta, CA, United States) were immediately placed into 1ml skim milk-tryptone-glucose-glycerol (STGG), transported at 4 °C to the laboratory within 2hours of collection and frozen at −80 °C for later batch culture. Swabs were inoculated onto Mannitol Salt Agar (MSA) (National Health Laboratory Services, Green Point Media Laboratory Cape Town, South Africa) and incubated at 35◦C for 18–24h in room temperature air prior to measurement. The presence or absence of the S *aureus*, S *pneumoniae*, H *influenzae* and M *catarrhalis* in culture results, as well as the date of the sample, the infant date of birth, and a few key clinical or demographic parameters including sex, site, and date of LRTI are included in the set of analysis measures.

Data analysis

The opportunity to work on a very densely sampled cohort provides different possible study designs. In the first instance infants will be categorised in two groups, those who developed pneumonia in the first year and those that did not, and the prevalence and pattern of colonisation of the four identified organisms will be described. Kaplan-Meier and time to event methods will be used to describe time to first organism acquisition.

To describe the pairwise temporal interactions Chi-squared and Fisher's exact tests will be used to estimate the odds ratio (OR) of positive-positive interaction between pairs of visits, versus no association or positive-negative interaction. This preliminary exploration will need to be adjusted for infant age in order to account for possible confounding.

Generalised estimating equations (GEE) or binomial models for repeated measures will be used to estimate the time varying association of organism colonisation with the development of pneumonia, while enabling adjustment by key confounders.

Multi-state models will be developed to estimate the probability of state transitions where states are defined by the presence or absence of specified organisms in a given sample.

Research feasibility

The DCHS parent study has been approved by all appropriate ethics boards and permission to use a sample from the data has been granted. My proposed supervisor has experience with the proposed methods and is a regular collaborator with the DCHS PIs (Prof H Zar and Prof M Nicol).

Description of risks and benefits

As this is a secondary analysis of data that has already been collected and de-identified, there are no direct risks or direct benefits to participants in this study. There is a small risk of loss of confidentiality due to the presence of personal data. There are anticipated general benefits as the work may help inform the pattern and role of key organisms in early life in association to lower respiratory tract infection.

Informed consent process

Written informed consent in the parent study was obtained from mothers and renewed annually. Additional consent was obtained from fathers, as identified by mothers, when possible. The DCHS received ethical approval from Human Research Ethics Committee (HREC), Faculty of Health Sciences University of Cape Town, Stellenbosch University and the Western Cape Provincial Research Committee.

Privacy and confidentiality

The privacy of participants is ensured in the analysis as the sample dataset has been de-identified by the DCHS data management team. No personal participant identifiers such as ID numbers or names are available to researchers. Study data will further be maintained on a password protected computer.

Dissemination of research findings

The findings will submitted as a mini-dissertation in partial fulfilment of the Masters of Public Health degree from the University of Cape Town and may also be submitted for publication in a peer reviewed journal.

Public health significance

In addition to the benefits alluded to in the problem identification section above, an understanding of underlying temporal organism interactions in the nasopharynx may help predict how novel vaccines against these pathogens may influence the ecology in this niche. It may also provide an avenue for design of more effective strategies to combating respiratory infections (Dunne et al. 2013).

References

Dunne, E. M., Smith-Vaughan, H. C., Robins-Browne, R. M., Mulholland, E. K., & Satzke, C. (2013). Nasopharyngeal microbial interactions in the era of pneumococcal conjugate vaccination. Vaccine, 31(19), 2333-2342.

Murphy, T. F., Bakaletz, L. O., & Smeesters, P. R. (2009). Microbial interactions in the respiratory tract. The Pediatric infectious disease journal, 28(10), S121-S126.

Xu, Q., Almudervar, A., Casey, J. R., & Pichichero, M. E. (2012). Nasopharyngeal bacterial interactions in children. Emerging infectious diseases, 18(11), 1738.

Zar, H. J., Barnett, W., Myer, L., Stein, D. J., & Nicol, M. P. (2014). Investigating the early-life determinants of illness in Africa: the Drakenstein Child Health Study. Thorax, thoraxjnl-2014.

Zar, H.J., Barnett, W., Stadler, A., Gardner-Lubbe, S., Myer, L., Nicol, M.P. (2016). Aetiology of childhood pneumonia in a well vaccinated South African birth cohort; a nested case-control study of the Drakenstein Child Health Study. Lancet Respiratory Medicine, 4(6):463-472.

Regev-Yochay G, Dagan R, Raz M, Carmeli Y, Shainberg B, Derazne E, et al. Association between carriage of Streptococcus pneumoniae and Staphylococcus aureus in Children. JAMA. 2004;292(6):716–20.

Bogaert D, Van Belkum A, Sluijter M, Luijendijk A, De Groot R, Rümke HC, et al. Colonisation by Streptococcus pneumoniae and Staphylococcus aureus in healthy children. Lancet. 2004;363(9424):1871–2.

Pettigrew MM, Gent JF, Revai K, Patel JA, Chonmaitree T. Microbial interactions during upper respiratory tract infections. Emerg Infect Dis. 2008;14(10):1584–91.

Jacoby, P., Watson, K., Bowman, J., Taylor, A., Riley, T. V., Smith, D. W., ... & Kalgoorlie Otitis Media Research Project Team. (2007). Modelling the co-occurrence of Streptococcus pneumoniae with other bacterial and viral pathogens in the upper respiratory tract. Vaccine, 25(13), 2458-2464.

Abdullahi, O., Karani, A., Tigoi, C. C., Mugo, D., Kungu, S., Wanjiru, E., ... & Scott, J. A. G. (2012). The prevalence and risk factors for pneumococcal colonization of the nasopharynx among children in Kilifi District, Kenya. PloS one, 7(2), e30787.

Lewnard, J. A., Givon-Lavi, N., Huppert, A., Pettigrew, M. M., Regev-Yochay, G., Dagan, R., & Weinberger, D. M. (2015). Epidemiological markers for interactions among Streptococcus pneumoniae, Haemophilus influenzae, and Staphylococcus aureus in upper respiratory tract carriage. The Journal of infectious diseases, 213(10), 1596-1605.

Anthony, L., Meehan, A., Amos, B., Mtove, G., Mjema, J., Malahiyo, R., ... & Reyburn, H. (2012). Nasopharyngeal carriage of Streptococcus pneumoniae: prevalence and risk factors in HIV-positive children in Tanzania. International Journal of Infectious Diseases, 16(10), e753-e757.

Rudan, I., O'brien, K. L., Nair, H., Liu, L., Theodoratou, E., Qazi, S., ... & Child Health Epidemiology Reference Group. (2013). Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. Journal of global health, 3(1).

Heinsbroek, E., Tafatatha, T., Chisambo, C., Phiri, A., Mwiba, O., Ngwira, B., ... & French, N. (2015). Pneumococcal acquisition among infants exposed to HIV in rural Malawi: A longitudinal household study. American journal of epidemiology, 183(1), 70-78.

Verani, J. R., Groome, M. J., Zar, H. J., Zell, E. R., Kapongo, C. N., Nzenze, S. A., ... & Madhi, S. A. (2016). Risk factors for presumed bacterial pneumonia among HIV-uninfected children hospitalized in Soweto, South Africa. The Pediatric infectious disease journal, 35(11), 1169-1174.

Madhi, S. A., Adrian, P., Kuwanda, L., Cutland, C., Albrich, W. C., & Klugman, K. P. (2007). Long-term effect of pneumococcal conjugate vaccine on nasopharyngeal colonization by Streptococcus pneumoniae—and associated interactions with Staphylococcus aureus and Haemophilus influenzae colonization—in HIV-infected and HIV-uninfected children. The Journal of infectious diseases, 196(11), 1662-1666.

Part B: Literature Review

Literature review objectives

The objectives of this literature review are to:

1. Describe the prevalence of nasopharyngeal carriage of four bacterial species (*Streptococcus pneumoniae, Staphylococcus aureus, Haemophilus influenzae* and *Moraxella catarrhalis)* in children in low- and middle-income countries, and explain any identified risk factors or confounding associations.
2. Identify and describe any organism interactions between the bacterial species.
3. Summarise statistical methods that have been applied previously to estimate either magnitude of, or direction of, organism interactions.

Search strategy

A search for relevant literature on organism interactions and nasopharyngeal carriage was conducted in PubMed in September 2018 using syntax constructed around the following search terms: upper respiratory tract; lower respiratory tract; Streptococcus pneumoniae; Haemophilus influenzae; Moraxella catarrhalis; Staphylococcus aureus; colonisation; children; interactions; pneumonia; nasopharyngeal colonisation.

Summary of the literature

The search resulted in 34 full papers which were reviewed. The search methodology is summarised in Supplement Figure 1 (Appendix) and further details are available in Supplement Table 1 (Appendix)

Bacterial nasopharyngeal carriage and risk factors

Variable nasopharyngeal (NP) carriage rates have been reported across different studies of varying populations, with the child's age appearing to modify carriage most strongly [1-13]. In a US cohort of children aged 6 - 36 months old [1], NP carriage prevalence per organism was 46% for *S. pneumoniae*, 7% for *S. aureus*, 32% for *H. influenzae* and 63% for *M. catarrhalis*. The authors of this study indicated that *M. catarrhalis* and *S. aureus* carriage was considerably below expected prevalence. Among Kenyan children between the ages of 12 – 59 months old, 50% of children who carried *S. pneumoniae* also co-carried *H. influenzae* [4]. Another Kenyan study which enrolled children between the ages of 3 - 59 months demonstrated the carriage rate for *S. pneumoniae* to be 66% [5]. A Gambian cohort of

children younger than 12 months of age (where NP samples were taken bi-weekly for the first six months of life and bi-monthly for remainder of the year) estimated that overall carriage rates of *S. pneumoniae, S. aureus, H. influenzae* and *M. catarrhalis* were 78%, 20%, 71% and 70%, respectively [6]. A similar study in Gambia [7] which enrolled children aged two months living in a peri-urban area, the carriage rates of *S. pneumoniae, S. aureus, H. influenzae* and *M. catarrhalis* were 62%, 50%, 30% and 32%, respectively. In Cape Town, South Africa among children who were characterised by poor socio-economic status, the overall carriage rate for *H. influenzae* was 46%, which was similar to the carriage rate estimated in children less than 12 months old at 44% [8]. A more recent study based in Cape Town on children with a median age of 36 months old and with suspected pulmonary Tuberculosis, the carriage rate of *S. pneumoniae, S. aureus* and *M. catarrhalis* were 42%, 22% and 64%, respectively [9]. Outside of sub-Saharan Africa, a Brazilian cohort of children 5 years or younger, of whom 53% had community acquired pneumonia the carriage rate of *S. pneumoniae* was 55% [10]. These data are summarised in Table 1.

Table 1. Bacterial species carriage rates in certain age groups and certain geographical regions

Bacteria prevalence	Study's geographic area and participants' age characteristics							
	USA[1] 6–12 months old	USA[2] 6-24 months old	Kenya[1] 12-59 months old	Kenya[2] 3-59 months old	Gambia[1] <12 months old	Gambia[2] 2 months old	South Africa[1] <12 months old	South Africa[2] median age 36 months
SP (%)	46%	30%	-	66%	78%	62%	-	42%
SA (%)	7%	-	-	-	20%	50%	-	22%
HI (%)	32%	12%	-	-	71%	30%	46%	-
MC (%)	63%	36%	-	-	70%	32%	-	64%
SP and HI co-carriage	-	-	50%	-	-	-	-	-

[1] and [2] indicates 2 different studies in the same country. SP – *S_pneumoniae;* SA - *S. aureus;* HI - *H. influenza*; MC - *M. catarrhalis*

A number of risk factors have been identified in children that are typically thought to be associated with organism carriage. These risk factors may include a social context, for example, day care attendance or the number of young children in a household. Clinical indicators may include a previous or current illness, as well as demographics, sex and age. In children aged 1 - 6 years old who had prior experience of penicillin-resistant *S. pneumoniae*

(PRSP) infection and who attended daycare, gender was not associated with any bacterial colonisation by either *S. pneumoniae, S. aureus, H. influenzae* or *M. catarrhalis*. Age, however was associated with colonisation by *M. catarrhalis* while cumulative time spent at the daycare centre was associated with colonisation by *H. influenzae* [11]. A similar lack of association between *S. pneumoniae* colonisation and gender was shown in children less than 5 years, in a Brazilian study [10]. This study showed that overcrowded housing was an important risk factor for *S. pneumoniae* colonisation especially in those individuals with a prior community acquired pneumonia. However, crowding, method of cooking and tobacco exposure were not associated with *S. pneumoniae, S. aureus, H. influenzae* or *M. catarrhalis* carriage in Venezuela, but were instead associated with poor nutrition [12]. In South African children hospitalised with Tuberculosis, those from a poor socio-economic background experienced higher rates of colonisation by *H. influenzae* compared to children from higher socio-economic backgrounds [11]. Among HIV-exposed neonates in Tanzania who had not received any pneumococcal conjugate vaccine (PCV), *S. aureus* colonisation was more prevalent in HIV-infected children compared to HIV-uninfected children. The converse applied for *S. pneumoniae* colonisation [15]. Moreover, the same study indicated that at 6 weeks of life colonisation by *S. aureus* was associated with urban area and having siblings younger than 10 years old. In contrast, at 6 months old only HIV infection remained associated with *S. aureus* colonisation [13]. This suggests that risk factors for *S. aureus* colonisation may be age dependent. Furthermore, at 3 months of age colonisation by *S. pneumoniae* and *H. influenzae* was associated with having siblings younger than 10 years old and colonisation by *M. catarrhalis* was associated with breast-feeding [13]. In preschool children aged 2 – 5 years old who had received antibiotics before undergoing an adenoidectomy for recurrent upper respiratory tract infections (URTIs), colonisation by at least one bacterial species (of *S. pneumoniae, S. aureus, H. influenzae* and *M. catarrhalis*) was associated with female gender and area of residence. Risk factors associated with individual bacterial colonisation included rural residence which was associated with a reduction in prevalence of *S. pneumoniae*; passive smoking and daycare attendance were associated with an increase in *H. influenzae*; female gender associated with an increase in *M. catarrhalis*; and residing in a rural area were associated with an increase in *S. aureus* [14].

Some studies have explored carriage rates specifically by disease status. Among children in South Africa younger than 5 years old, with Acute Otitis Media (AOM) the carriage rates of *S. pneumoniae, S. aureus, H. influenzae* and *M. catarrhalis* were 20%, 16%, 31% and 5%,

respectively [15]. These results were from cultured middle ear fluid samples rather than NP samples. In a much older cohort of children, median age 8.4 years old, who had a current lower respiratory tract infection (LRTI), NP carriage rates of *S. pneumoniae, S. aureus, H. influenzae* and *M. catarrhalis* were respectively 17%, 18%, 32% and 27% [16]. In a study of children 6 - 36 months old where NP samples were mostly collected within 7 days of an upper respiratory infection the overall carriage rates of *S. pneumoniae, S. aureus, H. influenzae* and *M. catarrhalis* were 46%, 7%, 32% and 63%, respectively [1]. The majority (96%) of these children had received at least one dose of PCV7 (vaccination available at time of research) at the time of study enrolment with as many as 95% not taking any antimicrobial drugs.

In Tanzania, HIV exposed neonates (six weeks old) of whom none had received any PCV, overall carriage rates of *S. pneumoniae, S. aureus, H. influenzae* and *M. catarrhalis* as 56%, 66%, 14% and 50%, respectively [13]. In healthy six month old children from USA [1], followed up until they were 24 months old, who had all received a dose of PCV-7 and other age appropriate vaccinations but no antimicrobial therapy, the carriage rates of *S. pneumoniae, H. influenzae* and *M. catarrhalis* were 30%, 12% and 36% respectively (see Table 1). Children with AOM had carriage rates of 53%, 48% and 43% respectively [17]. Moreover, the overall carriage rate by at least one of these bacterial species at healthy vs. at AOM visits were 57% vs. 90% [17] which suggests that the rate of NP bacterial colonisation can be expected to be higher in ill infants than in healthy children. These data are summarised in Table 2.

Table 2. Contrast of bacterial species carriage by infection and/or disease exposure

Bacteria specific prevalence	Illness and participants' age characteristics				
	AOM, <5years old, Middle ear fluid samples	LRTI, median age 8.4 years, NP samples	URTI 6-36 months old NP samples	HIV-exposed 6 weeks old, NP samples	AOM, 6-24 months old all received PCV7, NP samples
S. pneumoniae	20%	17%	46%	56%	53%
S. aureus	16%	18%	7%	66%	-
H. influenzae	31%	32%	32%	14%	48%
M. catarrhalis	5%	27%	63%	50%	43%

AOM: acute otitis media; LRTI: lower respiratory tract infection; URTI: upper respiratory tract infection, NP: nasopharyngeal

Healthy children may have a different set of risk factors compared to children with either a current or previous respiratory infection. In healthy Kenyan children of ages 3 – 59 months old who did not receive any pneumococcal vaccination, colonisation by *S. pneumoniae* was negatively associated with previous antibiotic use [5]. Meanwhile in Indonesian children aged 12 – 24 months old, the child's region of residence was associated with colonisation by *S. pneumoniae* and *M. catarrhalis* [17]. Furthermore, Israeli children <40 months old who did not receive any pneumococcal vaccination demonstrated that colonisation by *S. pneumoniae* was associated with their daycare attendance, having younger siblings and that they were older than 3 months; these three risk factors were all negatively associated with being colonised by *S. aureus* [18]. Although, maternal colonisation by *S. aureus* was associated with the child's subsequent colonisation by *S. aureus*, this was however not true for maternal carriage of S. *pneumoniae* and its subsequent colonisation in the child [18]. *Staphylococcus aureus* colonisation in children <6 months was not associated with either breastfeeding or the number of siblings [19].

Conjugate vaccine association with bacterial carriage before and after vaccination

A burden of disease related to vaccine-type *S. pneumoniae* persists in low- and middle-income countries (LMICs) despite certain serotypes of *S. pneumoniae* being effectively prevented by pneumococcal conjugate vaccine (PCV) PCV7, 9 and 13 [20]. The widespread introduction of PCVs raised some concerns over the possibility of a subsequent increase in *S. aureus* carriage and related disease post PCV vaccination [21]. This was due to the discovery of a negative association between *S. aureus* and *S. pneumoniae* [21]. Serotype replacement has been highlighted as one of the threats to conjugate vaccines, i.e. a biological phenomenon where NP colonisation by non-vaccine serotypes occurs post PCV vaccination. This risk might be heightened in neonates given the general consensus that NP flora develop in the first year of life and *S. aureus* carriage is typically highest among neonates. Risk generally declines with age. In South Africa, all children are expected to receive PCV and Hib vaccinations as part of the national immunisation programme. Hib is almost two decades old in the South African routine immunisation programme with PCV13 almost a decade old [22].

A number of studies have investigated the effects of the conjugate vaccines on NP carriage. In children aged 2 – 30 months, introduction to PCV7 did not affect *H. influenzae* or *S.*

aureus carriage rates [21]. A study compared children aged 12 – 59 months who received 2 doses of pneumococcal non-typeable *H. influenzae* protein-D conjugate vaccine (PHiD-CV) to children of the same age who received 2 doses of Hepatitis A vaccine, and showed that there was no significant increase in the carriage of NP bacteria such as *S. aureus* or *M. catarrhalis* due either to serotype replacement or the negative association between bacterial species [4]. Moreover, there was no increase in the carriage of *S. pneumoniae* or *H. influenzae* in either treatment group [4]. However, a different conclusion was reached by Van Gils et al. (2011) who showed that children who received a 2+1 dose of PCV (at the ages of 2, 4 and 11 months) there was evidence of temporary increase in *S. aureus* carriage at the age of 12 months while in the control treatment group (those who did not receive any PCV) there was a monotonic decrease in *S. aureus* carriage [23].

In Israel, children <40 months of age who had not been vaccinated with PCV, a negative association between vaccine-type strains of *S. pneumoniae* and colonisation by *S. aureus* was observed, additionally *S. aureus* carriage was higher in *S. pneumoniae* non-carriers compared to carriers and *S. pneumoniae* carriage was higher among *S. aureus* non-carriers compared to carriers [18]. Although a temporary increase in *S. aureus* was observed after adjusting for age [24], the authors argue that such an increase was not sufficient to support the hypothesis of an increase in *S. aureus* carriage post PCV vaccination due to the negative association between *S. aureus* and vaccine-type *S. pneumoniae*.

In South African children with AOM (3 months - 5 years) there was no significant difference in middle ear fluid carriage rates of *S. pneumoniae* and *H. influenzae* in those children who received a PCV7 dosage and those who did not [15]. Also, the number of PCV7 doses received by the child were not associated with nasal carriage of *S. pneumoniae*, vaccine-type and nonvaccine-type, or *S. aureus* in children aged between 6 weeks and 4 years [24]. Furthermore, even after adjusting for age, there was a temporary increase in *S. aureus* although this was found to be not sufficient to support the hypothesis that *S. aureus* carriage would increase after vaccination with PCV7 [24]..

Madhi et al. (2007) investigated the long term effects of PCV on NP colonisation by *S. pneumoniae* among both HIV-infected and uninfected South African children, and the interactions between pneumococcus species with other respiratory pathogens. Findings suggested that PCV9 and PCV13 increased NP colonisation by non-vaccine serotypes and reduced NP colonisation by vaccine serotypes [25]. Among South African children (mean

age 5.3 years) PCV was ineffective in reducing NP colonisation after 3 primary doses [22]. High *S. pneumoniae* and *H. influenzae* colonisation among HIV-infected South African children suggested that the burden of pneumococcal disease might be high among HIV infected children during periods of increased NP colonisation [25].

In a PCV7 dosing trial among 2 -30 months old children, the observed NP carriage rate of *S. pneumoniae* was 70% [21], while in children aged 6 weeks - 4 years old who attended daycare, of whom only a minority (27%) had received at least 1-dose of PCV7, the carriage rates for *S. pneumoniae* were 37% and 20% for *S. aureus* - however samples were nasal and not NP [24]. In contrast, children of mean age 5.3 years who had all received at least 1 dose of PCV9, the overall carriage rate of *S. pneumoniae, S. aureus* and *H. influenzae* were 54%, 35% and 57% respectively although carriage rates were significantly higher in HIV-infected children - 72% for *S. pneumoniae* and 74% for *H. influenzae* [25].These data are summarised in Table 3.

Table 3. PCV vaccination and subsequent carriage rate

Bacteria specific prevalence	Number of doses and study participants' age characteristics		
	PCV7 dosing trial, 2-30 months old	PCV7 - at least 1 dose 6 weeks – 4 years old	PCV9 -at least 1-dose mean age 5,3 years old
S. pneumoniae	70%	37%	54% (72% among HIV-infected infants)
S. aureus	-	20%	35%
H. influenzae	-	-	57% (74% among HIV-infected infants)
M. catarrhalis	-	-	-

Organism interactions

In Gambia among children <12 months old, positive interactions were observed between colonisation by *S. pneumoniae* with *H. influenzae* and *M. catarrhalis*; while a negative interaction between the colonisation by *S. pneumoniae* and *S. aureus* were observed [6]. Furthermore, it was noticed that there was an early colonisation by *H. influenzae* and *M. catarrhalis* which also frequently were co-carried alongside S. *pneumoniae*. In Sweden among children 1 – 6 years old, who experienced a penicillin-resistant *S. pneumoniae* there was a positive interaction between the colonisation of *S. pneumoniae* and *M. catarrhalis*

which decreased with age [11], while in this and a USA cohort, a negative interaction was observed between S. *pneumoniae* and *H. influenzae,* which became positive when *H. influenzae* co-colonised with *M. catarrhalis* [11, 1].

Meanwhile in Peruvian Andes children a positive interaction between colonisation by *S. pneumoniae* and *H. influenzae* and a negative interaction between colonisation by *S. pneumoniae* and *S. aureus* was seen [26]. A study distinguished between two levels of interactions [27], namely host (between subjects associations) and microbe (within subjects associations) and generally at both levels positive pairwise interactions were seen between *S. pneumoniae*, *H. influenzae* and *M. catarrhalis*.

Furthermore, among healthy Indonesian children of ages 12 – 24 months old, findings showed positive interactions between *S. pneumoniae* and *H. influenzae,* and between *H. influenzae* and *M. catarrhalis*; a negative interaction between *S. aureus* and *M. catarrhalis* was recorded [17]. In younger children, aged 6 – 12 months, positive interactions were observed between S. *pneumoniae* with *H. influenzae* and *M. catarrhalis,* and a negative interaction was observed between colonisation by S. *pneumoniae* and *S. aureus* [28].

Additionally among healthy 6 -24 months old children, a positive interaction between colonisation of S. *pneumoniae* and *M. catarrhalis*, and a negative interaction between colonisation by S. *pneumoniae* and *S. aureus* was recorded, while those study participants who acquired AOM demonstrated negative interactions between colonisation by S. *pneumoniae* and *H. influenzae* and between colonisation by *H. influenzae* and *M. catarrhalis* [15]. Preschool children 2 – 5 years old, who had received antibiotics before undergoing an adenoidectomy for upper respiratory tract infection (URTI), were found to have bacteria in the NP which were not associated with bacteria from the adenoidectomy surgical process, although in the adenoids there was a positive interaction between colonisation by *S. pneumoniae* and *H. influenzae* [16].

Furthermore, in children aged 6 weeks - 4 years, colonisation by *S. pneumoniae* (vaccine-type and non-vaccine-type) were negatively associated with colonisation by *S. aureus* [24]. Children (mean age 5.3 years)who had received at least 1 dose of PCV9 and who were also HIV-infected, showed that *S. aureus* interacted negatively with *S. pneumoniae* and *H. influenzae* [25]. *Streptococcus pneumoniae* interacted positively with colonisation of *H. influenzae* among both HIV-infected and HIV-uninfected children [25].

It has been suggested previously through a review of clinical studies that the community of bacteria in the upper respiratory tract is altered by respiratory viruses, and that respiratory viruses may promote bacterial colonisation of the lower respiratory tract [29]. It is generally understood that viral infections precede or predispose one to subsequent bacterial superinfections (or later co-infections) [30], given this, Murphy et al. (2009) reviewed studies of interactions between upper respiratory tract viruses and NP pathogens in relation to pathogenesis of upper respiratory tract infections (URTIs). These researchers observed synergism between respiratory viruses and NP bacteria. A study of Australian Aboriginal and non-Aboriginal children [27] looked at interactions between respiratory bacteria and rhinovirus and adenovirus both of which both are associated with otitis media (OM). At pathogen/microbe level, rhinovirus showed association with *S. pneumoniae, H influenzae and M catarrhalis* while adenovirus showed association only with *M. catarrhalis* colonisation [27].

Meanwhile among children of ages 6 – 12 months old, NP colonisation of *S. pneumoniae* was positively associated with human rhinovirus and enterovirus; colonisation by *H. influenzae* was positively associated with human rhinovirus and respiratory syncytial virus; colonisation by *M. catarrhalis* was positively associated with coronavirus and adenovirus; and colonisation by *S. aureus* was positively associated with influenza viruses [28]. In South African children with AOM aged 3 months - 5 years old, about 77% - 79% of children in the study had NP colonisation by at least one virus, accompanied by a co-carriage with either *S. pneumoniae, H. influenzae* or *M. catarrhalis* sampled from middle ear fluid. Additionally, colonisation by *S. aureus* was detected alongside at least one virus from the NP in 63% of the study participants [15].

Previously applied statistical methods

As the summary of relevant literature indicates, a majority of the studies investigated NP carriage or colonisation from cross-sectional data. Cross-sectional studies are typically limited in the statistical methods that can be applied to analyse longitudinal data, and are unable to ascertain the temporality of carriage and co-carriage events.

A review of the literature revealed that the most common method of analysis was logistic regression which was used to estimate pairwise bacterial interactions with outcome variable as binary bacterial colonisation status [17]. Additionally, interaction estimates between pairs of species were analysed by Chi-square or Fisher's exact odds ratios (ORs) computed

whenever statistically appropriate [31] with one bacterial species colonisation as the outcome variable and the other species as the exposure variable.

Interpretation of such an interaction is the same as the ordinary interpretation of the OR, with OR > 1 taken to be an indication of a positive interaction and OR < 1 to be an indication of a negative interaction [31, 32]. However, Suzuki et al. (date) highlighted concerns about an approach using ORs to indicate organism interactions, and they advise that potential differences in risks of enrolment for colonisation-positive and colonisation-negative individuals should be incorporated into the computation of such ORs, especially for case-control studies. They argued that unless a population based cross-sectional study is used, one could anticipate a different enrolment risk for sick or healthy children. A further criticism of this approach is that unadjusted ORs are insufficient since each pairwise comparison may depend on interactions with yet other organisms [32].

Longitudinal designs allow for other approaches. The most common approach to exploring bacterial interaction in longitudinal studies were population averaged or marginal models known as generalised estimating equations (GEE). GEEs have been used for fitting both univariate and multivariate logistic regression models with bacterial carriage coded as a binary outcome (presence/absence) [1, 21, 31]. The main difference in these studies were the choices of correlation structure and that of covariates for adjustment in the multivariable models. While GEE models allow for efficient adjustment for correlated data, they do lead to population averaged estimates of association which may not be desired in all contexts.

A few studies fit conditional multivariable mixed effects models [27] to estimate associations between pairs of bacterial species. These models fit two random components, distinguishing between microbe level correlations and host level correlations, where microbe level represents within-subject correlations and host level represents between-subjects correlations. It has been argued that such a model has an advantage over analysing longitudinal data cross-sectionally because it accounts for correlations in repeated observations from the same study participant and is also capable of differentiating between within-level and between-level correlations.

Another method identified in longitudinal studies that was used to understand bacterial colonisation in terms of acquisition, clearance and change of state hazard rates, was the Markov transition model [33]. Lipsitch et al.(2012) used such a model to understand transitions of multiple *S. pneumoniae* serotypes with acquisition defined as any transition

from state 0 (not colonised with *S. pneumoniae*) to state *i* (any of the other *S. pneumoniae* serotypes), and with transition specific hazard rate and clearance defined as any transition from state *i* (resident state) to state 0. Similarly change of serotype could be specified as any transition from state *i* to state *j* (challenge state). With such a model approximating n^2+n independent rates and to reduce dimensionality, the authors [33] decided to fit four separate models each estimating 2n acquisition and clearance rates but limiting the change of states. Limitations of the model were acknowledged as assumptions that transition hazards are constant for each of the transitions, and that homogeneity of hosts may not reflect reality, considering for example that age is known to be a strong influence for clearance. Their model also did not adjust for simultaneous carriage of multiple serotypes.

Finally, it is worth drawing attention to the possibility of measurement error which may depend on the method used to identify the organisms. Differences in yield of bacteria arising from detection by culture or by polymerase chain reaction (PCR) are expected, and a number of studies have contrasted the disparity in yield of bacteria from these methods. For instance, in a study of children in their first year of life, the carriage rates of *S. pneumoniae, S. aureus, H. influenzae* and *M. catarrhalis* were all much higher in PCR samples compared to biological culture samples [26, 34].

Discussion

Interactions between disease-causing organisms is important from a public health point of view. Although the seven studies I reviewed are not really enough for accurate country to country comparisons, there were indications that *S. pneumoniae* carriage or colonisation is typically high in all geographic areas varying between 30-78%. The carriage pattern for *S. pneumoniae* suggests its prevalence might be high in infants after which it declines to moderate colonisation when the child is 24 months (median age). *S. aureus* carriage patterns suggest that carriage is at least moderate in younger children i.e. 50% in 2 months old infants thereafter decreasing with age. The carriages patterns of *H. influenzae* and *M. catarrhalis* suggest moderately low carriage at younger age namely 12% for *H. influenzae* in children younger than 24 months of age (in USA) and 32% for *M. catarrhalis* among 2 months old infants (in Gambia) rising to approximately 70% in children <12 months old.

Data, as shown in Table 2 suggest that colonisation by *H. influenzae* is independent of respiratory infection although it is very low among HIV-exposed infants. Colonisation rates

of *S. pneumoniae* and *M. catarrhalis* are relatively higher in children with other concomitant health conditions such as upper respiratory infections and acute otitis (AOM) and in HIV-exposed infants, compared to those with lower respiratory tract infections and AOM (recovered from middle ear fluid samples rather than NP).

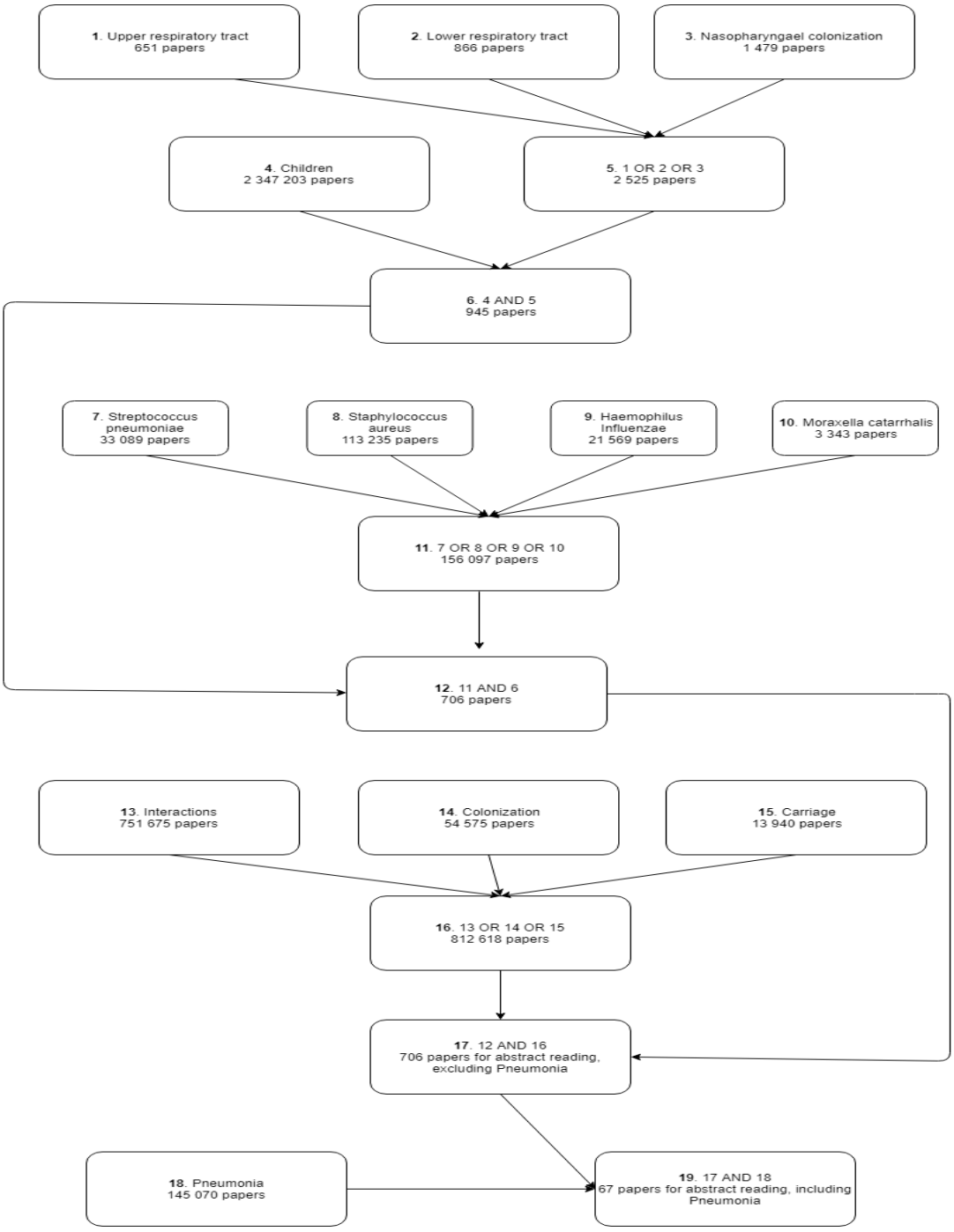
In Table 3, these data are not sufficient to support or refute the hypothesis of *S. aureus* increase post PCV vaccination by number of doses, however the PCV9 vaccine seem to be less effective among HIV-infected infants.

Positive interactions have been reported between *S. pneumoniae* and *M. catarrhalis,* while both positive and negative interactions have been reported with *H. influenzae* and *S. pneumoniae*. Positive interactions were recorded whenever *H. influenzae* and *M. catarrhalis* co-occurred. Among children with AOM, negative interactions have been described between *S. pneumoniae* colonisation wit*h H. influenzae* and between *H. influenzae* colonisation with *M. catarrhalis.* Meanwhile, negative interactions between colonisation by *S. aureus* with *S. pneumoniae* and *M. catarrhalis* have also been described. Moreover, among HIV-infected children *S. aureus* was also negatively associated with both *S. pneumoniae* and *H. influenzae,* while among HIV-infected and HIV-uninfected children *S. pneumoniae* and *H. influenzae* were both positively associated.

Cross-sectional methods used in longitudinal studies, for instance binning time into intervals, for calculations of odds ratios or interaction should be complemented by methods capable of controlling for a third variable for example other organisms interactions. Markov transition models seem ideal for investigating acquisition and clearance states and allow for the estimations for probability of change of states.

Appendix

Supplement Figure 1. Literature search strategy consort diagram



Literature Search Strategy.
Search terms: "Upper respiratory tract colonization", "Lower respiratory tract colonization", "Nasopharyngeal colonization", "Children", "Streptococcus pneumoniae", "Staphylococcus aureus", "Haemophilus influenzae", "Moraxella catarrhalis", "Interactions", "Colonization", "Carriage", "Pneumonia"

Supplement Table 1: Detailed search results

	Search term	**Date**	**Database**	**Results** (number of papers)
1	Upper respiratory tract	December 2018	PubMed	651
2	Lower respiratory tract	December 2018	PubMed	866
3	Nasopharyngeal colonisation	December 2018	PubMed	1 479
4	Children	December 2018	PubMed	2 347 203
5	1 or 2 or 3	December 2018	PubMed	2 525
6	4 and 5	December 2018	PubMed	945
7	Streptococcus pneumoniae	December 2018	PubMed	33 089
8	Staphylococcus aureus	December 2018	PubMed	113 235
9	Haemophilus influenzae	December 2018	PubMed	21 569
10	Moraxella catarrhalis	December 2018	PubMed	3 343
11	7 or 8 or 9 or 10	December 2018	PubMed	156 097
12	6 and 11	December 2018	PubMed	706
13	Interactions	December 2018	PubMed	751 675
14	Colonisation	December 2018	PubMed	54 575
15	Carriage	December 2018	PubMed	13 940
16	13 or 14 or 15	December 2018	PubMed	812 618
17	12 and 16	December 2018	PubMed	706
18	Pneumonia	December 2018	PubMed	145 070
19	17 and 18	December 2018	PubMed	67

References

[1] Pettigrew, M. M., Gent, J. F., Revai, K., Patel, J. A., & Chonmaitree, T. (2008). Microbial interactions during upper respiratory tract infections. *Emerging infectious diseases*, *14*(10), 1584.

[2] Dunne, E. M., Smith-Vaughan, H. C., Robins-Browne, R. M., Mulholland, E. K., & Satzke, C. (2013). Nasopharyngeal microbial interactions in the era of pneumococcal conjugate vaccination. *Vaccine*, *31*(19), 2333-2342.

[3] Lewnard, J. A., Givon-Lavi, N., Huppert, A., Pettigrew, M. M., Regev-Yochay, G., Dagan, R., & Weinberger, D. M. (2015). Epidemiological markers for interactions among Streptococcus pneumoniae, Haemophilus influenzae, and Staphylococcus aureus in upper respiratory tract carriage. *The Journal of infectious diseases*, *213*(10), 1596-1605.

[4] Feazel, L. M., Santorico, S. A., Robertson, C. E., Bashraheil, M., Scott, J. A. G., Frank, D. N., & Hammitt, L. L. (2015). Effects of vaccination with 10-valent pneumococcal non-typeable Haemophilus influenza protein D conjugate vaccine (PHiD-CV) on the nasopharyngeal microbiome of Kenyan toddlers. *PloS one*, *10*(6), e0128064.

[5] Abdullahi, O., Karani, A., Tigoi, C.C., Mugo, D., Kungu, S., Wanjiru, E., Jomo, J., Musyimi, R., Lipsitch, M. and Scott, J.A.G., 2012. The prevalence and risk factors for pneumococcal colonization of the nasopharynx among children in Kilifi District, Kenya. PloS one, 7(2), p.e30787.

[6] Kwambana, B. A., Barer, M. R., Bottomley, C., Adegbola, R. A., & Antonio, M. (2011). Early acquisition and high nasopharyngeal co-colonisation by Streptococcus pneumoniae and three respiratory pathogens amongst Gambian new-borns and infants. *BMC infectious diseases*, *11*(1), 175.

[7] Foster-Nyarko, E., Kwambana, B., Aderonke, O., Ceesay, F., Jarju, S., Bojang, A., McLellan, J., Jafali, J., Kampmann, B., Ota, M.O. and Adetifa, I., 2016. Associations between nasopharyngeal carriage of Group B Streptococcus and other respiratory pathogens during early infancy. BMC microbiology, 16(1), p.97.

[8] Hussey, G. D., Coetzee, G., Hitchcock, J., Van Schalkwyk, E., Van Wyk, H., & Kibel, M. (1994). Carriage of Haemophilus influenzae in Cape Town children. *South African Medical Journal*, *84*(3), 135-137.

[9] Dube, F. S., Kaba, M., Robberts, F. L., Tow, L. A., Lubbe, S., Zar, H. J., & Nicol, M. P. (2016). Respiratory microbes present in the nasopharynx of children hospitalised with suspected pulmonary tuberculosis in Cape Town, South Africa. *BMC infectious diseases*, *16*(1), 597.

[10] Rey LC, Wolf B, Moreira JL, Milatovic D, Verhoef J, Farhat CK. Antimicrobial susceptibility and serotypes of nasopharyngeal Streptococcus pneumoniae in children with pneumonia and in children attending day-care centres in Fortaleza, Brazil. Int J Antimicrob Agents. 2002 Aug 20(2):86-92. PubMed PMID: 12297356.

[11] Dahlblom, V., & Söderström, M. (2012). Bacterial interactions in the nasopharynx–effects of host factors in children attending day-care centers. *Journal of infection and public health*, *5*(2), 133-139.

[12] Verhagen, L. M., Hermsen, M., Rivera-Olivero, I. A., Sisco, M. C., de Jonge, M. I., Hermans, P. W., & de Waard, J. H. (2017). Nasopharyngeal carriage of respiratory pathogens in Warao Amerindians: significant relationship with stunting. *Tropical Medicine & International Health*, *22*(4), 407-414.

[13] Kinabo, G.D., van der Ven, A.J.A.M., Msuya, L.J., Shayo, A.M., Schimana, W., Ndaro, A., van Asten, H.A.G.H., Dolmans, W.M.V., Warris, A. and Hermans, P.W.M., 2013. Dynamics of nasopharyngeal bacterial colonisation in HIV-exposed young infants in T anzania. Tropical Medicine & International Health, 18(3), pp.286-295.

[14] Korona-Glowniak, I., Niedzielski, A., Kosikowska, U., Grzegorczyk, A., & Malm, A. (2015). Nasopharyngeal vs. adenoid cultures in children undergoing adenoidectomy: prevalence of bacterial pathogens, their interactions and risk factors. *Epidemiology & Infection*, *143*(4), 821-830

[15] Madhi, S.A., Govender, N., Dayal, K., Devadiga, R., Van Dyke, M.K., van Niekerk, N., Cutland, C.L., Adrian, P.V. and Nunes, M.C., 2015. Bacterial and respiratory viral interactions in the etiology of acute otitis media in HIV-infected and HIV-uninfected South African children. The Pediatric infectious disease journal, 34(7), p.753.

[16] Tenenbaum, T., Franz, A., Neuhausen, N., Willems, R., Brade, J., Schweitzer-Krantz, S., Adams, O., Schroten, H. and Henrich, B., 2012. Clinical characteristics of children with lower respiratory tract infections are dependent on the carriage of specific pathogens in the

nasopharynx. European journal of clinical microbiology & infectious diseases, 31(11), pp.3173-3182.

[17] Dunne, E.M., Murad, C., Sudigdoadi, S., Fadlyana, E., Tarigan, R., Indriyani, S.A.K., Pell, C.L., Watts, E., Satzke, C., Hinds, J. and Dewi, N.E., 2018. Carriage of Streptococcus pneumoniae, Haemophilus influenzae, Moraxella catarrhalis, and Staphylococcus aureus in Indonesian children: a cross-sectional study. PloS one, 13(4), p.e0195098.

[18] Regev-Yochay, G., Dagan, R., Raz, M., Carmeli, Y., Shainberg, B., Derazne, E., Rahav, G. and Rubinstein, E., 2004. Association between carriage of Streptococcus pneumoniae and Staphylococcus aureus in children. Jama, 292(6), pp.716-720.

[19] Peacock, S.J., Justice, A., Griffiths, D., De Silva, G.D.I., Kantzanou, M.N., Crook, D., Sleeman, K. and Day, N.P., 2003. Determinants of acquisition and carriage of Staphylococcus aureus in infancy. Journal of clinical microbiology, 41(12), pp.5718-5725.

[20] Clarke, E., Kampmann, B., & Goldblatt, D. (2016). Maternal and neonatal pneumococcal vaccination-where are we now? *Expert review of vaccines*, *15*(10), 1305-1317.

[21] Lewnard, J. A., Huppert, A., Givon-Lavi, N., Pettigrew, M. M., Regev-Yochay, G., Dagan, R., & Weinberger, D. M. (2016). Density, serotype diversity, and fitness of Streptococcus pneumoniae in upper respiratory tract cocolonization with nontypeable Haemophilus influenzae. *The Journal of infectious diseases*, *214*(9), 1411-1420.

[22] Verani, J.R., Groome, M.J., Zar, H.J., Zell, E.R., Kapongo, C.N., Nzenze, S.A., Mulligan, C., Moore, D.P., Whitney, C.G. and Madhi, S.A., 2016. Risk factors for presumed bacterial pneumonia among HIV-uninfected children hospitalized in Soweto, South Africa. The Pediatric infectious disease journal, 35(11), pp.1169-1174.

[23] Van Gils, E.J., Hak, E., Veenhoven, R.H., Rodenburg, G.D., Bogaert, D., Bruin, J.P., Van Alphen, L. and Sanders, E.A., 2011. Effect of seven-valent pneumococcal conjugate vaccine on Staphylococcus aureus colonisation in a randomised controlled trial. PloS one, 6(6), p.e20229.

[24] Dukers-Muijrers, N. H. T. M., Stobberingh, E., Beisser, P., Boesten, R. C. H., Jacobs, P., & Hoebe, C. J. P. A. (2013). Nasal carriage of Streptococcus pneumoniae serotypes and Staphylococcus aureus in Streptococcus pneumoniae-vaccinated and non-vaccinated young children. *Epidemiology & Infection*, *141*(3), 631-638.

[25] Madhi, S. A., Adrian, P., Kuwanda, L., Cutland, C., Albrich, W. C., & Klugman, K. P. (2007). Long-term effect of pneumococcal conjugate vaccine on nasopharyngeal colonization by Streptococcus pneumoniae and associated interactions with Staphylococcus aureus and Haemophilus influenzae colonization in HIV-infected and HIV-uninfected children. *The Journal of infectious diseases*, *196*(11), 1662-1666.

[26] Grijalva, C.G., Griffin, M.R., Edwards, K.M., Williams, J.V., Gil, A.I., Verastegui, H., Hartinger, S.M., Vidal, J.E., Klugman, K.P. and Lanata, C.F., 2013. Cohort profile: the study of respiratory pathogens in Andean children. International journal of epidemiology, 43(4), pp.1021-1030.

[27] Jacoby, P., Watson, K., Bowman, J., Taylor, A., Riley, T.V., Smith, D.W., Lehmann, D. and Kalgoorlie Otitis Media Research Project Team, 2007. Modelling the co-occurrence of Streptococcus pneumoniae with other bacterial and viral pathogens in the upper respiratory tract. Vaccine, 25(13), pp.2458-2464.

[28] van den Bergh, M.R., Biesbroek, G., Rossen, J.W., de Steenhuijsen Piters, W.A., Bosch, A.A., van Gils, E.J., Wang, X., Boonacker, C.W., Veenhoven, R.H., Bruin, J.P. and Bogaert, D., 2012. Associations between pathogens in the upper respiratory tract of young children: interplay between viruses and bacteria. PloS one, 7(10), p.e47711.

[29] Lee, K. H., Gordon, A., & Foxman, B. (2016). The role of respiratory viruses in the etiology of bacterial pneumonia: an ecological perspective. *Evolution, medicine, and public health*, *2016*(1), 95-109.

[30] Murphy, T. F., Bakaletz, L. O., & Smeesters, P. R. (2009). Microbial interactions in the respiratory tract. The Pediatric infectious disease journal, 28(10), S121-S126.

[31] Xu, Q., Almudervar, A., Casey, J. R., & Pichichero, M. E. (2012). Nasopharyngeal bacterial interactions in children. Emerging infectious diseases, 18(11), 1738.

[33] Suzuki, M., Dhoubhadel, B. G., Yoshida, L. M., & Ariyoshi, K. (2014). Nasopharyngeal bacterial interactions in children. *Emerging infectious diseases*, *20*(2), 323-324.

[33] Lipsitch, M., Abdullahi, O., D'Amour, A., Xie, W., Weinberger, D. M., Tchetgen, E. T., & Scott, J. A. G. (2012). Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. *Epidemiology (Cambridge, Mass.)*, *23*(4), 510.

[34] Tsai, M.H., Huang, S.H., Chen, C.L., Chiu, C.Y., Hua, M.C., Liao, S.L., Yao, T.C., Lai, S.H., Yeh, K.W., Wang, M.P. and Huang, J.L., 2015. Pathogenic bacterial nasopharyngeal colonization and its impact on respiratory diseases in the first year of life: the PATCH Birth Cohort Study. The Pediatric infectious disease journal, 34(6), pp.652-658.

Part C: Manuscript

Title: Temporal interactions of microbiota in longitudinal nasopharyngeal samples

Author: Brian Rambau^

Affiliation: Division of Epidemiology and Biostatistics, School of Public Health and Family Medicine, University of Cape Town, South Africa

Target journal: International Journal of Epidemiology*

Word count and references:
Abstract: 230 (Max: 250)
Main text: 2596 (Max: 3000)
References: 15

^ Per University of Cape Town MPH guidelines, sole authorship of this manuscript is reported.

* This manuscript is prepared in accordance with the guidelines set out by the Annals of Epidemiology, except that tables and figures have been included in-line for ease of reading, and page numbering is left contiguous with the rest of the document.

Abstract

Background In the aetiology of bacterial pneumonia it is generally accepted that infection is preceded by nasopharyngeal (NP) colonisation. Several studies have shown that NP flora develop and establish in first year of life or early in childhood. Temporal organism interactions have rarely been explored due to the need for longitudinal data and the complexity of the statistical models for recurrent events.

Methods We sought to identify, describe, and quantify the temporal interactions existing between selected key bacteria colonising the nasopharynx in infants (up to 1 year of life). We then compared patterns in those children who go on to develop pneumonia versus those who do not. We applied multi-state models to examine the patterns of transition among states of colonisation while accounting for key confounders.

Results There were 760 individuals included in the analysis, with a total of 16,346 NP samples available and a median 364 person-days (IQR 346 – 365 person-days). There were temporally sustained positive interactions between *Streptococcus pneumoniae* with *Haemophilus influenzae; Streptococcus pneumoniae* with *Moraxella catarrhalis;* and *Haemophilus influenzae* with *Moraxella catarrhalis*. Moreover, the extent of association generally decreased with child age. *Staphylococcus aureus* had consistent negative interactions with other organisms.

Conclusions We have developed a flexible framework that characterised interactions between organisms while accounting for key time-varying confounders. We have demonstrated that multi-state models provide a useful approach to supplement classical modelling approaches.

Introduction

In the aetiology of bacterial pneumonia it is generally accepted that infection is preceded by nasopharyngeal (NP) colonisation [1]. Several studies have shown that NP bacterial flora develop and establish in the first year of life or early in childhood. Of the pathogens that colonise the nasopharynx the most commonly studied are *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Moraxella catarrhalis* and *Staphylococcus aureus*, in part because they are known contributors to lower respiratory tract infections [2].

Colonisation of the nasopharynx of young children is often asymptomatic [3], and concurrent with complex organism interactions. Colonisation is impacted upon by different risk factors such as antimicrobial treatment and/or vaccination [2], as well as environmental exposures, which may act as mediators.

A body of literature exists on organism interactions [1, 3-7].Temporal organism interactions have however not been explored in depth, especially regarding whether observed colonisation patterns are resultant of the host's immunity, synergism or antagonism between bacterial species interactions. An understanding of underlying temporal organism interactions in the nasopharynx may help predict how novel vaccines against these bacterial species may influence the ecology in this niche area.

We sought to identify, describe, and quantify temporal interactions existing between selected key bacteria present in the nasopharynx in infants (up to 1 year of life), and to compare these patterns in children who go on to develop pneumonia versus those who do not. We also explored methods for estimation of temporal interactions. We investigated patterns of carriage and co-carriage in a cohort of South African children residing in a peri-urban area near Cape Town, South Africa. We describe multivariable marginal models and multi-state models appropriate for the analysis of longitudinal data.

Participants and Methods

Study design

Data from the Drakenstein Child Health Study (DCHS) [15, 16] are used. DCHS is a population-based birth cohort study investigating early-life determinants of child health in a peri-urban setting in the Western Cape province of South Africa. The cohort includes more than 1000 mother-infant pairs engaged with the public sector health services from antenatal care until at least five years of age. Follow-up is ongoing- all infants have completed at least two years of follow up.

For this study, data come from nasopharyngeal (NP) swab samples taken every two weeks for the first year of life in a subset of more than 800 infants whose mothers enrolled in monitored particular cohort of the DCHS. In addition to routine mother-infant visits, study personnel actively surveyed episodes of illness, including lower respiratory tract infections (LRTI).

Data collection method

NP swabs were taken every two weeks and at LRTI episodes. Collected NP swabs (Copan flocked swab, FLOQSwabsTM, COPAN Diagnostics, Murrieta, CA, United States) were immediately placed into 1ml skim milk-tryptone-glucose-glycerol (STGG), transported at 4 °C to the laboratory within 2 hours of collection and frozen at −80 °C for later batch culture. Swabs were inoculated onto Mannitol Salt Agar (MSA) at the National Health Laboratory Services, Green Point Media Laboratory Cape Town, South Africa and incubated at 35 °C for 18–24 hours prior to measurement. The presence or absence of four organisms: *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Moraxella catarrhalis* and *Staphylococcus aureus* was recorded. Demographic and other data were collected by trained study personnel using electronic case report forms.

Data measures

Clinical and demographic parameters included sex of the child, date of birth, HIV status of mother and infant, and residential location. If an episode of lower respiratory tract infection occurred in the first year of life these data were recorded. Age groupings were used to stratify results for descriptive purposes and these included: birth-4 weeks, 5-12 weeks, 13-24 weeks and 25-52 weeks. These age categories were selected as they represent common stages in infant grown. All intervals are inclusive.

Carriage was defined on a per-sample and per-individual basis. Individuals were defined with positive carriage if they had any positive sample in the follow-up period. The frequency (percentage) of positive samples/individuals was calculated taking as the denominator the total samples available overall or in the specified time period. Co-carriage required the same sample to be a positive sample for both specified organisms.

Statistical analysis

Total person-days were calculated as the age (in days) at the last available sample for each individual. Age (in days) at the first positive sample was used to determine the age of first carriage. The cumulative time spent positive for carriage/colonisation was calculated as the sum of the intervals between consecutive positive samples (for each organism) including the following interval up to the first negative sample. Time regarded as negative was taken to be the individual total person-days minus the cumulative time regarded as positive.

Temporal pairs are defined as a pair of consecutive samples from the same individual. Clearance events are defined in a temporal pair, where the first sample is +ve (carriage) for a given organism and the second sample is –ve (no carriage). Acquisition events are the reverse (-ve → +ve). Steady state events are those remaining in a positive carriage state (+ve → +ve) or those remaining in a non-carriage state (-ve → -ve) over the temporal pair.

Generalised mixed effects models using a logit link were used to estimate the time varying association of organism carriage with the development of pneumonia, while enabling adjustment by key confounders. Crude and age-adjusted odds ratio (OR) and 95% confidence intervals were estimated for the positive-positive temporal interaction versus no association or positive-negative interaction.

Kaplan-Meier estimates were used to reflect time to first carriage. Markov multi state models (MSMs) were used to describe how organism acquisition in infancy rolls out with age and associated probabilities of clearance, new acquisition or persistent carriage. We fitted piecewise constant MSMs with knots fit at 28days (d), 84d, 168d and 365d corresponding to the *a priori* defined age categories. All MSMs were adjusted for age at sample collection and the allowable state-transitions are described in Supplement Figure 1. Observed versus expected prevalence plots were used to graphically assess model fit. Estimated transition probabilities were summarised numerically and graphically.

Ethical approval

All study procedures were carried out under approval by the University of Cape Town Human Research Ethics Committee for the DCHS study (HREC: 401/2009) and for this analysis (HREC: 240/2018). Participants in the original study provided written consent at recruitment and consent was updated annually if maternal participants wished to remain in the study.

Results

There were 760 individuals included in the analysis, with a total of 16,346 NP samples available and a median 364 person-days (IQR 346 – 365 person-days). *S. pneumoniae* was predominant, found in >55% of all samples and demonstrating carriage in >95% of individuals at least once by 12 months of age (Figures 1 and 2, Supplementary Table 1). *H. influenzae* and *M. catarrhalis* had 35% and 41% positive carriage respectively in all samples, and appeared at least once in >90% of individuals. *S. aureus* was less common (25% of samples and 88% of individuals) and also had a strikingly different pattern of first acquisition compared to the other three organisms (Figures 1 and 2), demonstrating a rapid increase in carriage prevalence until approximately 4 weeks and subsequently declining. *S. aureus* carriage appeared earlier than other bacterial species (Figure 2; Table 1) with first carriage event occurring at a median 28 days of age (IQR: 15 – 58 days) as compared to *H. influenzae* where first carriage occurred at a median 97 days of age (IQR: 57 – 155 days). Individuals remained positive for *S. aureus* for the shortest cumulative time, accruing a median 121 days (IQR: 55 – 180 days) and median 36% of the total person-days while both *H. influenzae* and *S. pneumoniae* had positive carriage for >200 days on average (Table 1).

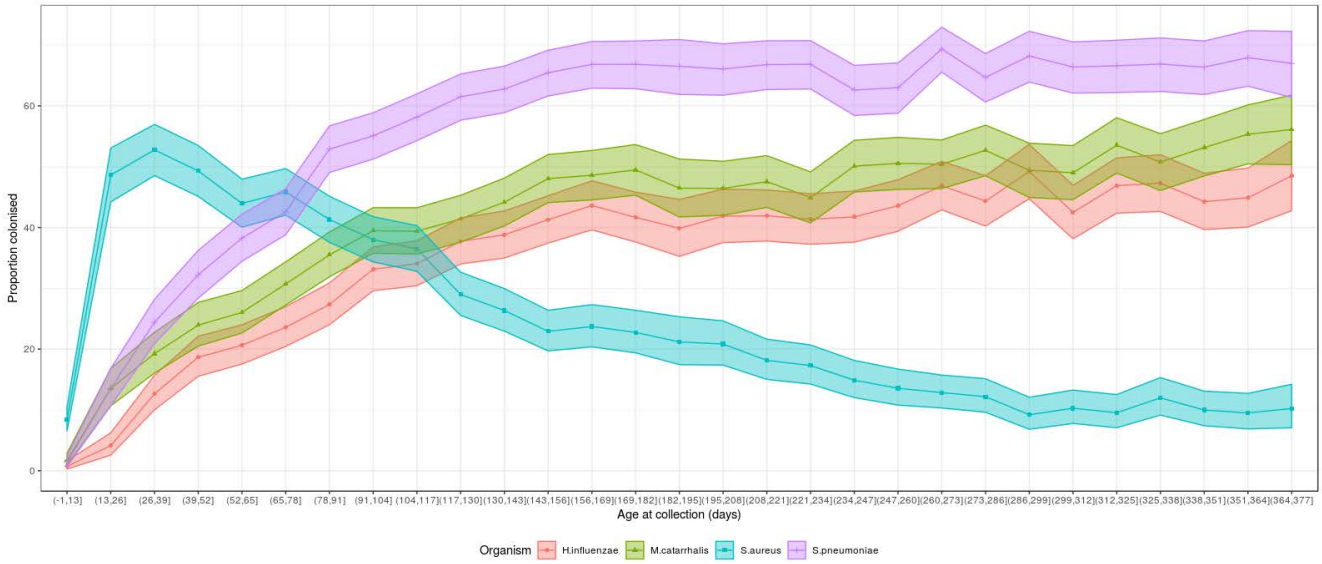


Figure 1. Prevalence (pointwise 95% CI) of sample carriage by each bacterial organism over time.

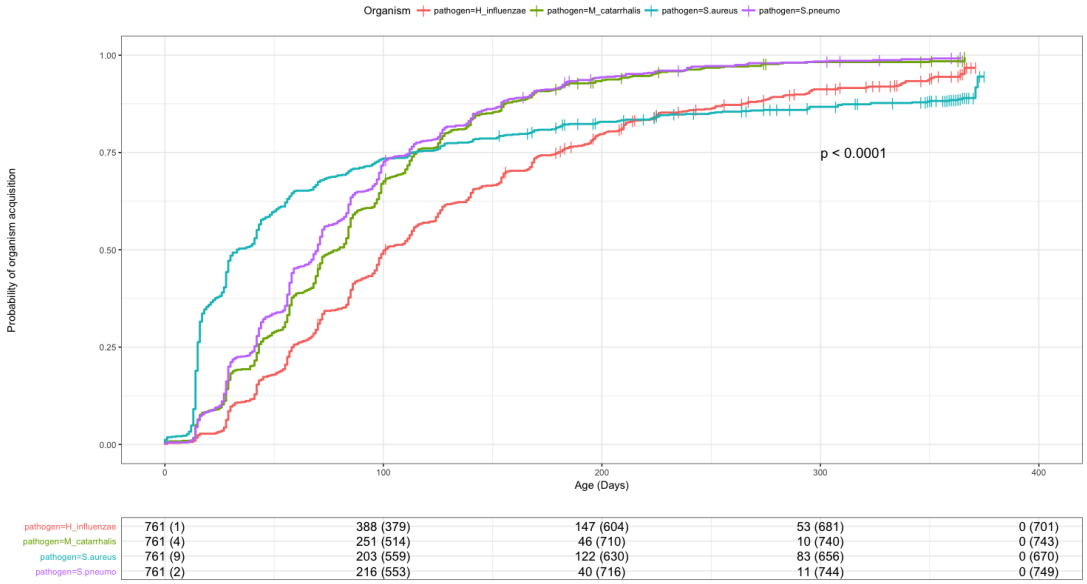


Figure 2. Cumulative probability of first carriage of bacterial organism by time (age in days).

Table 1. Statistical summary of estimated time spent with organism carriage (sum of person-days between consecutive positive samples for each individual) and the time spent without carriage (time spent positive subtracted from the total person-days for each individual)

Organism	Age (days) first carriage# Median (IQR)	Cumulative time (days) spent positive^ Median (IQR)	Percent person time positive Median (IQR)	Time (days) spent negative* Median (IQR)	Median (IQR) person-days
S. aureus	28 (15, 68)	121 (55, 180)	35.9% (16.0, 54.0)	219 (152, 286)	364 (346, 365)
S. pneumoniae	69 (41, 104)	246 (168, 301)	73.2% (52.9, 86.5)	88 (45, 156)	364 (346, 365)
H. influenzae	97 (57, 155)	168 (86, 238)	49.7% (26.7, 68.7)	169 (99, 239)	364 (346, 365)
M. catarrhalis	73 (43, 112)	200 (140, 264)	60.8% (43.9, 74.5)	127 (84, 193)	364 (346, 365)

among those ever positive; ^ Calculated as the median (IQR) of the individual sum of person-days between consecutive colonized samples including the right hand period (up to the first non-colonised sample). *Calculated as the median (IQR) of the individual person-days minus the individual days spent colonised

S. pneumoniae had the highest co-carriage prevalence with *H. influenzae* and *M. catarrhalis* (both 25%) but this varied by age category (Table 2). *H. influenzae* also co-occurred frequently with *M. catarrhalis* (19%). In contrast, *S. aureus* co-carriage was least prevalent with *S. pneumoniae* (12%), *H. influenzae* (5%) or *M. catarrhalis* (6%). Co-carriage frequencies differed considerably by age category, at least partially reflecting the relative prevalence of carriage by age. The co-carriage frequency of *H. influenzae* and *M. catarrhalis;* *H. influenzae* and *S. pneumoniae;* and *S. pneumoniae* and *M. catarrhalis*, increased with age,

from a rate of 4-6% in the first four weeks of life to 26-34% in the second 6 months of life. Co-carriage of *S. aureus* with all other organisms peaked between 5 and 12 weeks of age. Carriage and co-carriage rates were similar among those children who experienced LRTI compared to those who did not (Table 2). Seasonal carriage varied (Supplementary Figure 2), but to a lesser extent as compared to variance by age.

Overall, *M. catarrhalis* demonstrated the most stochastic or short term carriage over time, with 16% of temporal pairs representing clearance (carriage → absent) and 18% of temporal pairs being acquisition events (absent → carriage). This contrasts with *S. aureus* which had clearance events occurring in only 10% of temporal pairs and a similar rate for acquisition events (Supplementary Table 2).

S. aureus was negatively associated with all three bacterial species *S. pneumoniae* OR (95% CI): 0.68 (0.63 – 0.73), *H. influenzae* 0.41 (0.38 – 0.45) and *M. catarrhalis* 0.41 (0.38 – 0.45) in NP samples. Positive association interactions were evident between *S. pneumoniae* and *H. influenzae* OR (95% CI): 3.18 (2.97 – 3.41) and *S. pneumoniae* and *M. catarrhalis* 2.18 (2.04 – 2.32). *H. influenzae* and *M. catarrhalis* were also positively associated OR (95% CI): 2.58 (2.41 – 2.76).

Table 2. Percentage (count) of samples co-occurring by age category and LRTI status

Among all children							
Age category	**N**	**SA + SP**	**SA + HI**	**SA + MC**	**SP + HI**	**SP + MC**	**HI + MC**
Birth – 4 weeks	3019	10.2% (307)	3.88% (117)	5.3% (160)	5.17% (156)	6.36% (192)	4.21% (127)
5 – 12 weeks	2768	20.7% (572)	8.24% (228)	11.1% (307)	20.1% (557)	22% (609)	15% (416)
13 – 24 weeks	4416	14.4% (634)	6.5% (287)	6.75% (298)	30.8% (1361)	32.1% (1417)	23.2% (1024)
25-52 weeks	6143	6.85% (421)	3.48% (214)	4.49% (276)	33.7% (2068)	35.6% (2186)	26% (1598)
Birth – 1 year (overall)	16346	11.8% (1934)	5.18% (846)	6.37% (1041)	25.3% (4142)	26.9% (4404)	19.4% (3165)
Among children who experienced LRTI							
Age category	**N**	**SA + SP**	**SA + HI**	**SA + MC**	**SP + HI**	**SP + MC**	**HI + MC**
Birth – 4 weeks	1379	9.2% (127)	3.7% (51)	5.1% (70)	5.3% (73)	7.2% (100)	4.1% (56)
5 – 12 weeks	1307	20.6% (270)	7.5% (98)	10.9% (142)	17.8% (233)	21.3% (279)	15.1% (197)
13 – 24 weeks	2051	15.6% (319)	7.3% (149)	7.6% (155)	28.8% (590)	31.7% (650)	23.2% (476)
25-52 weeks	2838	7.3% (208)	3.3% (93)	4.3% (123)	32.7% (927)	36.2% (1026)	26.0% (738)
Birth – 1 year (overall)	7575	12.2% (924)	5.2% (391)	6.5% (490)	24.1% (1823)	27.1% (2055)	19.4% (1467)
Among children who did not experience LRTI							
Age category	**N**	**SA + SP**	**SA + HI**	**SA + MC**	**SP + HI**	**SP + MC**	**HI + MC**
Birth – 4 weeks	1640	11.0% (180)	4.0% (66)	5.5% (90)	5.1% (83)	5.6% (92)	4.3% (71)
5 – 12 weeks	1461	20.7% (302)	8.9% (130)	11.3% (165)	22.2% (324)	22.6% (330)	15.0% (219)
13 – 24 weeks	2365	13.3% (315)	5.8% (138)	6.0% (143)	32.6% (771)	32.4% (767)	23.2% (548)
25-52 weeks	3305	6.4% (213)	3.7% (121)	4.6% (153)	34.5% (1141)	35.1% (1160)	26.0% (860)
Birth – 1 year (overall)	8771	11.5% (1010)	5.2% (455)	6.3% (551)	26.4% (2319)	26.8% (2349)	19.4% (1698)

SA: *S aureus*; SP: *S pneumoniae;* HI: *H influenza*; MC: *M catarrhalis;* LRTI: lower respiratory tract infection

Mixed effects models adjusting for sex, site, season of birth, and age found temporally sustained positive associations between the co-carriages of *S. pneumoniae* with both *H. influenzae,* and *M. catarrhalis,* but no association with *S. aureus*. Moreover the extent generally decreased with infant's age (Table 3). Models including an indicator variable for LRTI in the first year of life had similar associations between organisms, and no evidence of meaningful associations with LRTI (Table 3).

Table 3. Associations with *S. pneumoniae* by generalised mixed effects models with a logit link and random effect term for individual. Each organism represents an independent model.

Models without LRTI status	**Main effects models^**		**Age interaction models^**	
	Est (SE)	**P-value**	**Est (SE)**	**P-value**
S. aureus	0.024 (0.05)	0.617	0.021 (0.05)	0.683
S. aureus * age interaction	-	-	-0.008 (0.05)	0.874
H. influenza	0.73 (0.04)	< 0.0001	0.86 (0.04)	< 0.0001
H. influenza * age interaction	-	-	-0.55 (005)	< 0.0001
M. catarrhalis	0.52 (0.04)	< 0.0001	0.54 (0.04)	< 0.0001
M. catarrhalis * age interaction	-	-	-0.51 (0.04)	< 0.0001
Models including LRTI status				
S. aureus	0.024 (0.05)	0.619	0.28 (0.07)	< 0.0001
S. aureus * age interaction	-	-	-0.009 (0.05)	0.864
LRTI in the first year of life	0.13 (0.10)	0.279	0.13 (0.12)	0.269
H. influenza	0.73 (0.04)	< 0.0001	0.79 (0.04)	< 0.0001
H. influenza * age interaction	-	-	-0.55 (0.04)	< 0.0001
LRTI in the first year of life	-0.01 (0.12)	0.934	-0.043 (0.12)	0.721
M. catarrhalis	0.52 (0.04)	< 0.0001	0.54 (0.04)	< 0.0001
M. catarrhalis * age interaction	-	-	-0.51 (0.04)	< 0.0001
LRTI in the first year of life	0.12 (0.12)	0.319	0.09 (0.12)	0.456

^ Adjusted for sex, site, season of birth, scaled age at collection, main effect for age retained in model and statistically significant for all models ($p < 0.0001$);

Multi-state models were built to estimate transition probabilities between co-carriage states, as well as estimate the association between predefined risk factors and transition between states. The general multi-state model scheme can be found in Supplement Figure 1. The model allowed for four possible states: no carriage, carriage with two organisms, or carriage with one organism and negative for the other. The multi-state process was assumed to be Markovian, with transition intensities modelled as piecewise constant.

The probability of acquisition of *S. pneumoniae* is modified by earlier carriage of *H. influenzae* or *M. catarrhalis*, however the reverse is not true. Positive *H. influenzae* carriage increases the probability of acquisition of *S. pneumoniae* with transition probabilities from 0.15 (95% CI 0.14-0.17) to 0.36 (95% CI 0.17, 0.54) for age categories over 28 days (Table 4), compared to the probability of acquisition of *S. pneumoniae* alone at 0.015 (95% CI 0.043-0.076) to 0.088 (95% CI 0.075-0.10) over the same age categories. This is in contrast to the probability of *H. influenzae* acquisition, which is similar with and without prior S. *pneumoniae* carriage. Similarly, carriage of *M. catarrhalis* increases the risk of *S. pneumoniae* acquisition (probability 0.24 vs 0.14 between 84 -168d of age), but *S. pneumoniae* carriage does little to modify the risk of *M. catarrhalis* acquisition. There was no evidence of differences between the acquisition of *S. pneumoniae* or *S. aureus* with relation to co-carriage.

There is no difference in the clearance of *S. pneumoniae* related to *H. influenzae* carriage, yet clearance of *H. influenzae* before 6 months of age is far less likely if coming from a state of co-carriage (probability between 0.04 - 0.07) compared to sole carriage (probability 0.23 - 0.12). The only evidence of differences in clearance probability in the models investigating *S. pneumoniae* and *M. catarrhalis* are in the probability of *M. catarrhalis* clearance before 28 days, which is 0.24 (95% CI 0.15 - 0.38) if carried alone and only 0.058 (55% CI 0.01 - 0.30) if carried with S. *pneumoniae,* though these confidence intervals overlap. Multi-state models

incorporating LRTI as a fifth absorbing state did not converge, nor did the secondary analysis of presented models on data subsets by LRTI status. Diagnostics for the presented models were accurate based on visual inspection of the predicted vs observed probability plots.

Table 4A. Predicted transition probabilities (95% confidence intervals) for the probability of state changes in each time period estimated with pairwise multistate models including S. *pneumoniae*, adjusted for sex, site, season of birth.

	Predicted probability (95% CI) of transition in the next 14 days with covariates at their means			
S. *pneumoniae with H. influenzae*	**0-28d**	**28-84d**	**84-168d**	**365d +**
Acquisition events				
Gain SP (from negative SP/HI)	0.019 (0.01, 0.034)	0.051 (0.043, 0.076)	0.066 (0.054, 0.080)	0.088 (0.075, 0.10)
Gain SP (from positive HI only)	0.062 (0.00, 0.13)	0.15 (0.14, 0.17)	0.21 (0.19, 0.22)	0.36 (0.17, 0.54)
Gain HI (from negative SP/HI)	0.067 (0.052, 0.086)	0.11 (0.098, 0.18)	0.13 (0.12, 0.15)	0.15 (0.13, 0.16)
Gain HI (from positive SP only)	0.011 (0.00, 0.35)	0.13 (0.095, 0.32)	0.19 (0.15, 0.22)	0.21 (0.19, 0.22)
Gain both (from negative SP/HI)	0.017 (0.00, 0.030)	0.013 (0.011, 0.78)	0.036 (0.029, 0.057)	0.050 (0.043, 0.069)
Clearance events				
Lose SP (from SP positive only)	0.37 (0.18, 0.63)	0.20 (0.00, 0.24)	0.19 (0.16, 0.23)	0.19 (0.17, 0.22)
Lose SP (from positive both)	0.35 (0.061, 0.66)	0.17 (0.13, 0.22)	0.15 (0.13, 0.17)	0.20 (0.18, 0.21)
Lose HI (from HI positive only)	0.23 (0.15, 0.38)	0.12 (0.00, 0.14)	0.087 (0.075, 0.10)	0.086 (0.076, 0.097)
Lose HI (from positive both)	0.044 (0.00, 0.57)	0.065 (0.043, 0.095)	0.063 (0.051, 0.078)	0.068 (0.060, 0.076)
Lose both	0.088 (0.048, 0.90)	0.051 (0.00, 0.089)	0.039 (0.030, 0.056)	0.043 (0.037, 0.053)
S. *pneumoniae with M. catarrhalis*	**0-28d**	**28-84d**	**84-168d**	**168+**
Acquisition events				
Gain SP (from negative SP/MC)	0.066 (0.051, 0.081)	0.088 (0.059, 0.10)	0.14 (0.12, 0.16)	0.16 (0.14, 0.18)
Gain SP (from positive MC only)	0.081 (0.00, 0.16)	0.18 (0.082, 0.28)	0.24 (0.15, 0.27)	0.31 (0.24, 0.33)
Gain MC (from negative SP/MC)	0.065 (0.052, 0.084)	0.10 (0.089, 0.29)	0.12 (0.11, 0.15)	0.12 (0.11, 0.15)
Gain MC (from positive SP only)	0.13 (0.01, 0.19)	0.10 (0.048,0.34)	0.14 (0.092, 0.17)	0.15 (0.12, 0.18)
Gain both (from negative SP/MC)	0.011 (0.00, 0.067)	0.020 (0.014, 0.69)	0.042 (0.033 ,0.096)	0.064 (0.050, 0.20)
Clearance events				
Lose SP (from SP positive only)	0.28 (0.17, 0.46)	0.26 (0.00, 0.33)	0.21 (0.18, 0.27)	0.19 (0.16, 0.23)
Lose SP (from positive both)	0.25 (0.074, 0.42)	0.26 (0.17, 0.32)	0.26 (0.20, 0.30)	0.29 (0.24, 0.31)
Lose MC (from MC positive only)	0.24 (0.15, 0.38)	0.099 (0.00, 0.20)	0.082 (0.070, 0.17)	0.090 (0.076, 0.17)
Lose MC (from positive both)	0.058 (0.01, 0.30)	0.068 (0.045, 0.097)	0.060 (0.049, 0.090)	0.063 (0.056, 0.090)
Lose both (from positive both)	0.060 (0.033, 0.86)	0.045 (0.00, 0.64)	0.030 (0.024, 0.45)	0.034 (0.030, 0.32)

Table 4B. Predicted transition probabilities (95% confidence intervals) for the probability of state changes in each time period estimated with pairwise multistate models including S. *pneumoniae* and *S aureus,* adjusted for sex, site, season of birth.

S. *pneumoniae with S aureus*	**0-28d**	**28-84d**	**84+ d**
Acquisition events			
Gain SP (from negative SP/SA)	0.22 (0.19, 0.25)	0.092 (0.077, 0.11)	0.059 (0.053, 0.067)
Gain SP (from positive SA only)	0.15 (0.00, 0.24)	0.12 (0.10, 0.16)	0.052 (0.00, 0.063)
Gain SA (from negative SP/SA)	0.057 (0.043, 0.074)	0.11 (0.091, 0.13)	0.19 (0.18, 0.21)
Gain SA (from positive SP only)	0.090 (0.00, 0.12)	0.11 (0.093, 0.14)	0.13 (0.00,0.16)
Gain both(from negative SP/SA)	0.020 (0.00, 0.073)	0.016 (0.013, 0.13)	0.013 (0.00, 0.17)
Clearance events			
Lose SP (from SP positive only)	0.097 (0.073, 0.21)	0.12 (0.099,0.14)	0.20 (0.17, 0.34)
Lose SP (from positive both)	0.063 (0.024, 0.19)	0.14 (0.11, 0.17)	0.29 (0.19, 0.31)
Lose SA (from SA positive only)	0.14 (0.070, 0.39)	0.11 (0.085,0.14)	0.097 (0.085, 0.15)
Lose SA (from positive both)	0.23 (0.15, 0.42)	0.081 (0.061,0.11)	0.077 (0.059, 0.090)
Lose both	0.028 (0.017, 0.74)	0.039 (0.024, 0.094)	0.031 (0.028, 0.74)

Discussion

Statistical modelling of temporal interactions may identify important organism level associations that change over time or vary in subgroups of children. We found positive and sustained interactions between *S. pneumonia* and both *H. influenzae* and *M. catarrhalis,* where our models indicated that preceding carriage or colonisation with either *H. influenzae* or *M. catarrhalis* may increase the risk of colonisation with *S. pneumonia.*

Timing of carriage and overall prevalence of carriage are in line with other findings in similar populations [1-7] with overall high exposure to *S. pneumonia*, *H. influenza* and *M. catarrhalis* during the first year of life and rapid and early exposure to *S. aureus.* Co-carriage is common, occurring in 5% - 26% of all samples, depending on the age of the children and the specific organisms. Similarly, clearance and acquisition events made up approximately 20% of all transitions suggesting that patterns of acquisition and clearance are dynamic. Carriage, co-carriage and transition frequency did not vary appreciably when comparing children who experienced LRTI in the first year of life versus those who had no lower tract respiratory infections. This suggests that the underlying pattern of exposure may be similar in

those children who do not experience LRTI in the first year of life, and so further investigation into the timing of exposure in the period before LRTI onset is critical.

The predicted transition probabilities from the multi-state models showed evidence of changing temporal risk of acquisition and clearance of *S. pneumonia* depending on co-carriage with other organisms.

This study has several strengths including an unprecedented sample frame of repeated NP samples in over 700 children during the first year of life and which accounted for children who experienced or did not experience lower respiratory tract infections (LRTIs). This represented a rare opportunity to carefully identify and analyse temporal interactions between organisms.

As an observational cohort, the positive associations that we found do not necessarily represent causal effects, and even though the temporal order was accounted for, organism carriage is not the same as colonisation or infection. Treatment with antibiotics and the impact of vaccination have not been accounted for here, and findings may change when considering colonisation with organisms rather than simple carriage, as well as adjustment for known environmental and clinical risk factors. The cohort is situated in a specific context and setting of South Africa, and so the results found here may not be generalizable to other settings.

We have presented an initial investigation into temporal interactions in a densely sampled longitudinal data set. Further analyses might consider different definitions of carriage, to get closer to the idea of colonisation, and account for washout periods due to antibiotics and/or other treatments. The multi-state models applied here have demonstrated their use as an important statistical tool in understanding longitudinal data and change between states over time.

Conclusion

Application of multi-state models to the temporal changes in nasopharyngeal carriage of four key disease-causing bacteria, has provided insights into the organism interactions occurring in the host. This work demonstrates that multi-state models are a useful additional statistical tool in cases where there are sufficient frequencies of longitudinal measures and repeat events.

Funding

BR was supported during his studies by the Medical Research Council of South Africa in terms of the National Health Scholars Programme from funds provided for this purpose by the National Department of Health/Public Health Enhancement Fund and by the South African Department of Science and Technology/National Research Foundation (DST-NRF), Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, Stellenbosch, South Africa.

References

[1] Dunne, E. M., Smith-Vaughan, H. C., Robins-Browne, R. M., Mulholland, E. K., & Satzke, C. (2013). Nasopharyngeal microbial interactions in the era of pneumococcal conjugate vaccination. Vaccine, 31(19), 2333-2342.

[2] Rudan, I., O'brien, K.L., Nair, H., Liu, L., Theodoratou, E., Qazi, S., Lukšić, I., Walker, C.L.F., Black, R.E., Campbell, H. and Child Health Epidemiology Reference Group, 2013. Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. Journal of global health, 3(1).

[3] Bogaert D, Van Belkum A, Sluijter M, Luijendijk A, De Groot R, Rümke HC, et al. Colonisation by Streptococcus pneumoniae and Staphylococcus aureus in healthy children. Lancet. 2004;363(9424):1871–2.

[4] Abdullahi, O., Karani, A., Tigoi, C.C., Mugo, D., Kungu, S., Wanjiru, E., Jomo, J., Musyimi, R., Lipsitch, M. and Scott, J.A.G., 2012. The prevalence and risk factors for pneumococcal colonization of the nasopharynx among children in Kilifi District, Kenya. PloS one, 7(2), p.e30787.

[5] Anthony, L., Meehan, A., Amos, B., Mtove, G., Mjema, J., Malahiyo, R., Yin, J.K., Oftadeh, S., Gilbert, G.L., Shingadia, D. and Reyburn, H., 2012. Nasopharyngeal carriage of Streptococcus pneumoniae: prevalence and risk factors in HIV-positive children in Tanzania. International Journal of Infectious Diseases, 16(10), pp.e753-e757.

[6] Heinsbroek, E., Tafatatha, T., Chisambo, C., Phiri, A., Mwiba, O., Ngwira, B., Crampin, A.C., Read, J.M. and French, N., 2015. Pneumococcal acquisition among infants exposed to HIV in rural Malawi: a longitudinal household study. American journal of epidemiology, 183(1), pp.70-78.

[7] Verani, J.R., Groome, M.J., Zar, H.J., Zell, E.R., Kapongo, C.N., Nzenze, S.A., Mulligan, C., Moore, D.P., Whitney, C.G. and Madhi, S.A., 2016. Risk factors for presumed bacterial pneumonia among HIV-uninfected children hospitalized in Soweto, South Africa. The Pediatric infectious disease journal, 35(11), pp.1169-1174.

[8] Murphy, T. F., Bakaletz, L. O., & Smeesters, P. R. (2009). Microbial interactions in the respiratory tract. The Pediatric infectious disease journal, 28(10), S121-S126.

[9] Xu, Q., Almudervar, A., Casey, J. R., & Pichichero, M. E. (2012). Nasopharyngeal bacterial interactions in children. Emerging infectious diseases, 18(11), 1738.

[10] Pettigrew MM, Gent JF, Revai K, Patel JA, Chonmaitree T. Microbial interactions during upper respiratory tract infections. Emerg Infect Dis. 2008;14(10):1584–91.

[11] Jacoby, P., Watson, K., Bowman, J., Taylor, A., Riley, T.V., Smith, D.W., Lehmann, D. and Kalgoorlie Otitis Media Research Project Team, 2007. Modelling the co-occurrence of Streptococcus pneumoniae with other bacterial and viral pathogens in the upper respiratory tract. Vaccine, 25(13), pp.2458-2464.

[12] Madhi, S.A., Adrian, P., Kuwanda, L., Cutland, C., Albrich, W.C. and Klugman, K.P., 2007. Long-term effect of pneumococcal conjugate vaccine on nasopharyngeal colonization

by Streptococcus pneumoniae—and associated interactions with Staphylococcus aureus and Haemophilus influenzae colonization—in HIV-infected and HIV-uninfected children. The Journal of infectious diseases, 196(11), pp.1662-1666.

[13] Regev-Yochay G, Dagan R, Raz M, Carmeli Y, Shainberg B, Derazne E, et al. Association between carriage of Streptococcus pneumoniae and Staphylococcus aureus in Children. JAMA. 2004;292(6):716–20.

[14] Lewnard, J. A., Givon-Lavi, N., Huppert, A., Pettigrew, M. M., Regev-Yochay, G., Dagan, R., & Weinberger, D. M. (2015). Epidemiological markers for interactions among Streptococcus pneumoniae, Haemophilus influenzae, and Staphylococcus aureus in upper respiratory tract carriage. The Journal of infectious diseases, 213(10), 1596-1605.

[15] Zar, H. J., Barnett, W., Myer, L., Stein, D. J., & Nicol, M. P. (2014). Investigating the early-life determinants of illness in Africa: the Drakenstein Child Health Study. Thorax, thoraxjnl-2014.

[16] Zar, H.J., Barnett, W., Stadler, A., Gardner-Lubbe, S., Myer, L., Nicol, M.P. (2016). Aetiology of childhood pneumonia in a well vaccinated South African birth cohort; a nested case-control study of the Drakenstein Child Health Study. Lancet Respiratory Medicine, 4(6):463-472.

Manuscript supplement

Supplement Table 1: Percentage (frequency) of samples and individuals with organism carriage in age groups at the time of collection. Individuals were considered positive if they had any positive sample collected during the relevant age category.

Full cohort

		Samples					Individuals			
Age category	N samples	*SA*	*SP*	*HI*	*MC*	N	*SA*	*SP*	*HI*	*MC*
Birth – 4 weeks	3019	37.8% (1140)	20.4% (617)	10.7% (323)	15.7% (475)	730	65.9% (481)	45.9% (335)	26.3% (192)	39.9% (291)
5 – 12 weeks	2768	40.5% (1121)	51.9% (1436)	29.3% (811)	35.9% (994)	732	63.9% (468)	73.4% (537)	52.5% (384)	70.2% (514)
13 – 24 weeks	4416	23.7% (1048)	65.2% (2881)	40.9% (1805)	46.4% (2048)	734	56.0% (411)	91.1% (669)	76.3% (560)	87.7% (644)
25-52 weeks	6143	12.0% (737)	66.3% (4070)	44.8% (2752)	51% (3135)	684	49.4% (338)	94.0% (643)	82.5% (564)	91.1% (623)
Birth – 1 year (total)	16346	24.8% (4046)	55.1% (9004)	34.8% (5691)	40.7% (6652)	738	88.0% (650)	98.6% (728)	89.2% (685)	98.5% (727)

Among children who experienced LRTI

		Samples					Individuals			
Age category	**N samples**	***SA***	***SP***	***HI***	***MC***	**N**	***SA***	***SP***	***HI***	***MC***
Birth – 4 weeks	1379	36.3% (501)	19.8% (273)	10.2% (141)	16.9% (233)	318	65.0% (215)	45.6% (151)	27.2% (90)	42.6% (141)
5 – 12 weeks	1307	39.7% (519)	50.0% (640)	27.8% (364)	35.6% (466)	337	61.1% (204)	71.6% (239)	52.1% (174)	70.4% (235)
13 – 24 weeks	2051	25.7% (527)	63.0% (1293)	40.7% (835)	47.9% (982)	334	59.1% (199)	91.4% (308)	76.6 % (258)	89.0% (300)
25-52 weeks	2838	11.7% (333)	65.9% (1871)	44.0% (1298)	51.6% (1466)	331	49.7% (158)	94.7% (301)	85.5% (272)	90.3% (287)
Birth – 1 year (total)	7575	24.8% (1880)	53.8% (4077)	34.1% (2588)	41.5% (3175)					

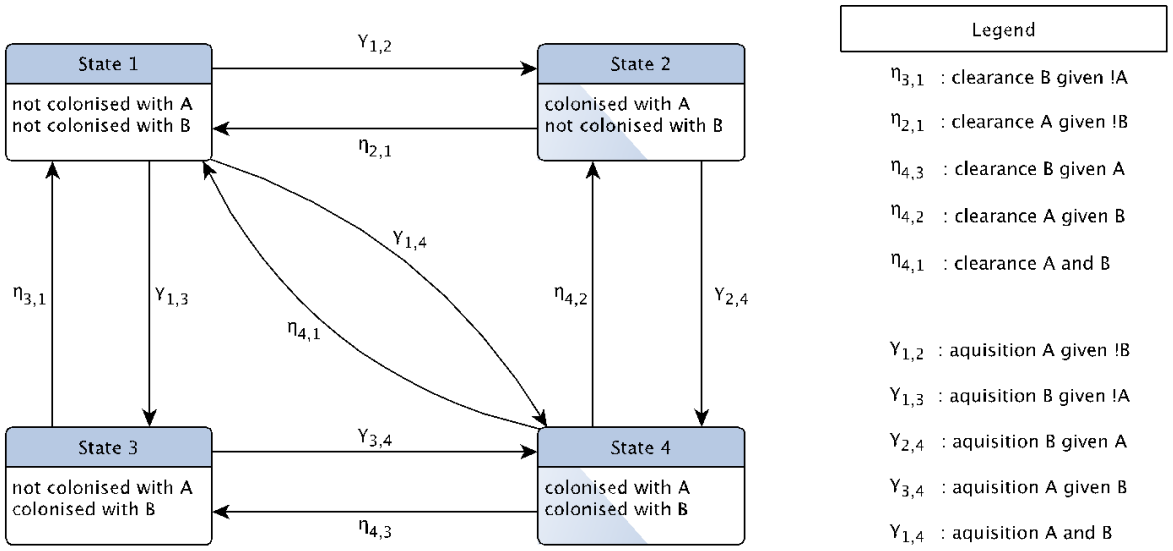
Among children who never experienced LRTI

		Samples					Individuals			
Age category	**N samples**	***SA***	***SP***	***HI***	***MC***	**N**	***SA***	***SP***	***HI***	***MC***
Birth – 4 weeks	1640	39.0% (639)	21.0% (344)	11.1% (182)	14.8% (242)	399	66.7% (266)	46.1% (184)	25.6% (102)	37.6% (150)
5 – 12 weeks	1461	41.2% (602)	54.5% (796)	30.6% (447)	36.1% (528)	398	66.3% (264)	74.9% (298)	52.8% (210)	70.1% (279)
13 – 24 weeks	2365	22.0% (521)	67.1% (1588)	41.1% (970)	45.1% (1066)	397	53.4% (212)	90.9% (361)	76.1% (302)	86.6% (344)
25-52 weeks	3305	12.2% (404)	66.5% (2199)	45.5% (1504)	50.5% (1669)	366	49.2% (180)	93.4% (342)	79.8% (292)	91.8% (336)
Birth – 1 year (total)	8771	24.7% (2166)	56.2% (4927)	35.4% (3103)	40.0% (3505)					

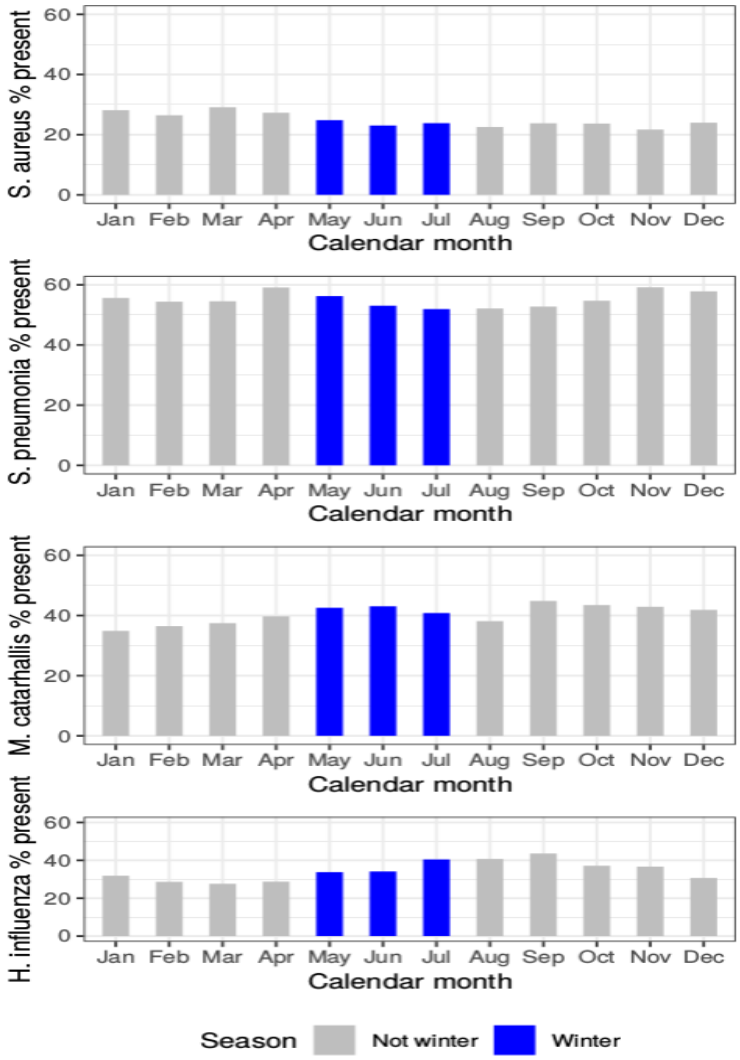
SA: *S aureus*; SP: *S pneumoniae;* HI: *H influenza*; MC: *M catarrhalis;* LRTI: lower respiratory tract infection

Supplement Table 2. Percentage (frequency) of each pair of sequential samples (n = 15,608 pairs) that had a clearance, acquisition or remained in existing state from observation *t* to observation *t+1*.

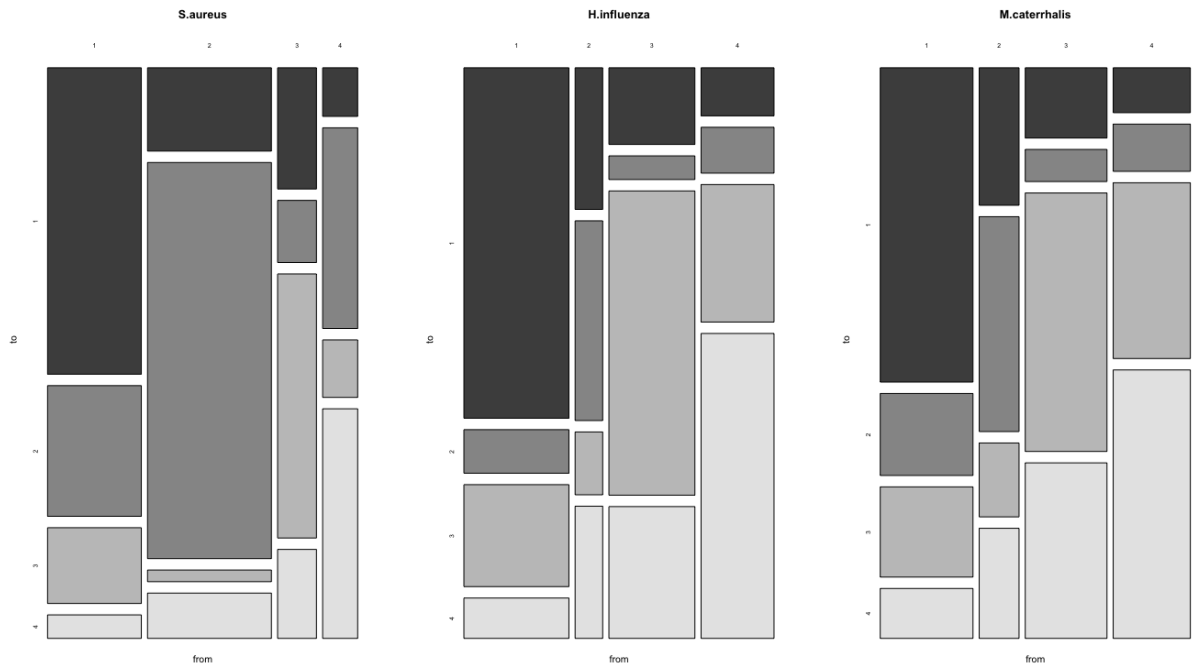
	S. aureus	***S. pneumoniae***	***H. influenzae***	***M. catarrhalis***
Clearance **+ve → -ve**	9.7% (1583)	9.5% (1545)	11.7% (3468)	15.6% (2541)
Acquisition **-ve → +ve**	10.0% (1633)	12.4% (2029)	13.6% (2214)	17.9% (2915)
No change **-ve → -ve**	61.5% (10019)	31.1% (5071)	49.2% (8012)	39.4% (6427)
No change **+ve → +ve**	14.6% (2373)	42.7% (6963)	21.3% (3468)	22.8% (3725)



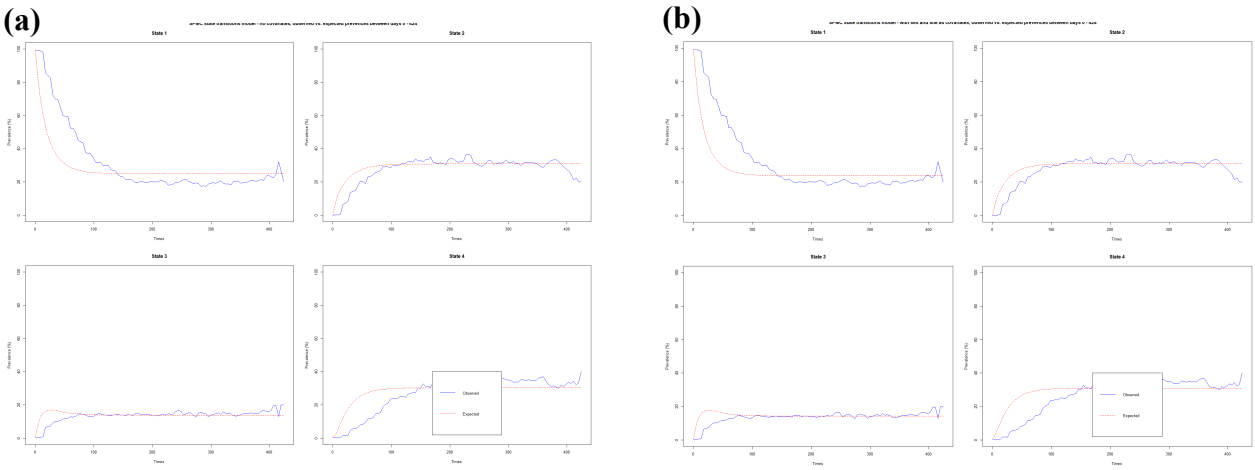
Supplement Figure 1. Schematic of two-organism states and transitions. Some organism pairs had additional transitions non-allowable due to small numbers.



Supplement Figure 2. Seasonal carriage by organism and calendar month.



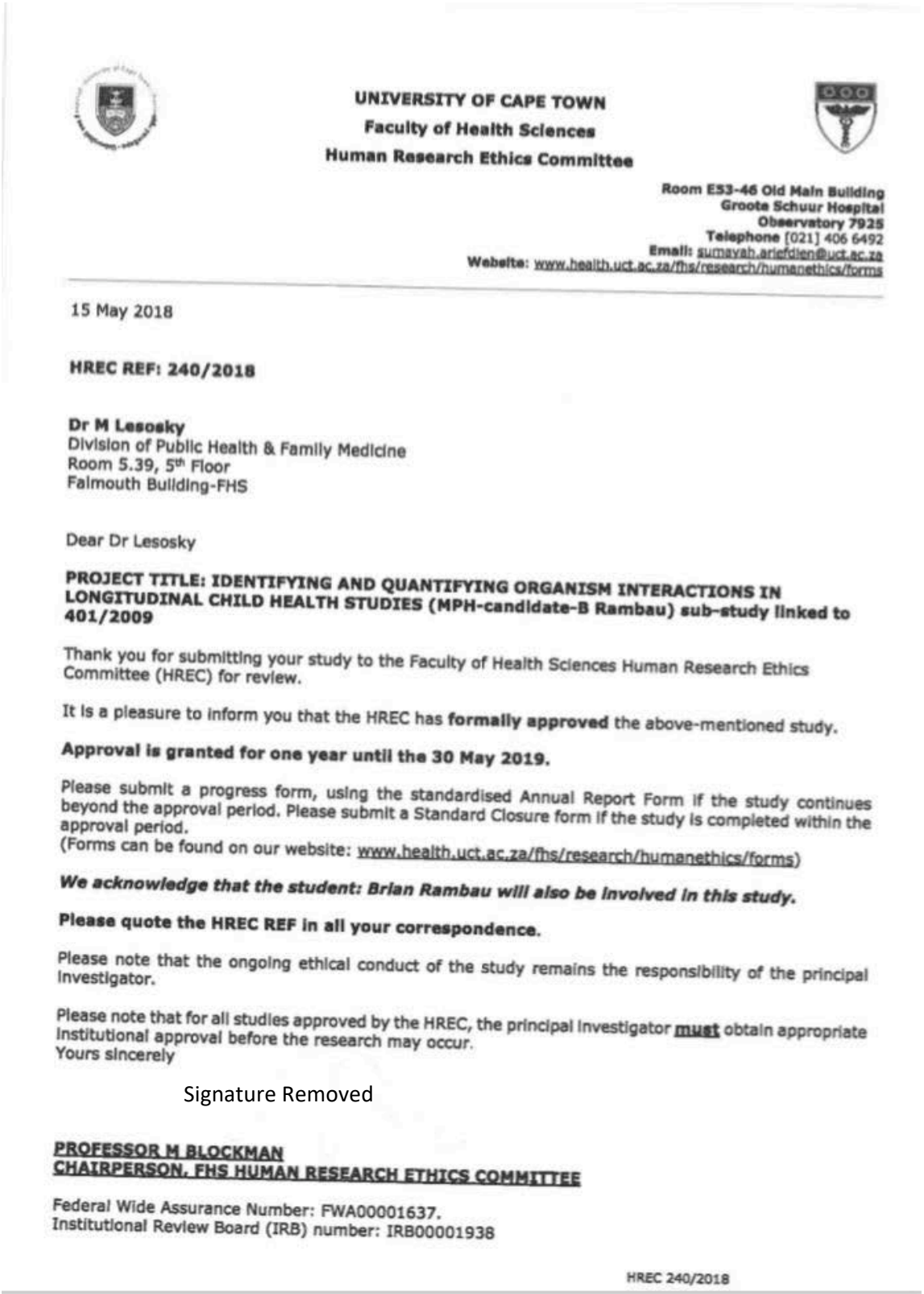
Supplement Figure 3. Mosaic plot of co-colonisation transition matrix for *S. pneumonia* with other organisms. Rows represent origin state (time *t*), columns represent next state (time *t+1*) and size reflects probability of transition.



Supplement Figure 4. Selected model diagnostics for multi-state models baseline (a) and covariate adjusted (b) for the *S. pneumonia* and *M. catarrhalis* model.

Part D: Appendices

Appendix 1: HREC approval letters



UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E53-46 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6492
Email: sumayah.ariefdien@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

15 May 2018

HREC REF: 240/2018

Dr M Lesosky
Division of Public Health & Family Medicine
Room 5.39, 5th Floor
Falmouth Building-FHS

Dear Dr Lesosky

PROJECT TITLE: IDENTIFYING AND QUANTIFYING ORGANISM INTERACTIONS IN LONGITUDINAL CHILD HEALTH STUDIES (MPH-candidate-B Rambau) sub-study linked to 401/2009

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee (HREC) for review.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 30 May 2019.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

We acknowledge that the student: Brian Rambau will also be involved in this study.

Please quote the HREC REF in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate institutional approval before the research may occur.
Yours sincerely

Signature Removed

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE

Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938

HREC 240/2018

UNIVERSITY OF CAPE TOWN

FACULTY OF HEALTH SCIENCES
Human Research Ethics Committee

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938) This serves as notification of annual approval, including any documentation described below.			
☑ Approved	Annual progress report	Approved until/next renewal date	30/08/19
☐ Not approved	See attached comments		
Signature Chairperson of the HREC	Signature Removed pp	Date Signed	12/09/2018

Comments to PI from the HREC

Principal Investigator to complete the following:

1. Protocol Information

Date (when submitting this form)	22 Aug 2018		
HREC REF Number	401/2009	Current Ethics Approval was granted until	30 Aug 2018
Protocol title	Drakenstein Child Health Study		
Protocol number (if applicable)	Protocol v1.15		
Are there any sub-studies linked to this study?	✓ Yes ☐ No		
If yes, could you please provide the HREC Ref's for all sub-studies? **Note:** A separate FHS016 must be submitted for each sub-study.	Details included in appendix.		
Principal Investigator	Prof Heather Zar		
Department / Office Internal Mail Address	Department of Paediatric and Child Health Red Cross War Memorial Children's Hospital		

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

Appendix 2: Instructions for Authors –Annals of Epidemiology

ANNALS OF EPIDEMIOLOGY

TABLE OF CONTENTS

AUTHOR INFORMATION PACK

ISSN: 1047-2797

Annals of Epidemiology is a peer reviewed, international journal devoted to **epidemiologic research** and methodological development. The journal emphasizes the application of epidemiologic methods to issues that affect the **distribution** and **determinants** of **human illness** in diverse contexts. Its primary

focus is on **chronic** and **acute** conditions of diverse etiologies and of major importance to clinical medicine, public health, and health care delivery.

Annals encourages the use of **epidemiology** in a multidisciplinary approach to understanding **disease etiology**. Review articles, reports from U.S. Federal and International sources, Editorials, Commentaries, Brief Communications, Letters to the Editor, Book Reviews, and selected papers from major symposia are also published.

For more information concerning The American College of Epidemiology (ACE), please visit their home page atacepidemiology.org .

AUDIENCE

Epidemiologists, public health workers, clinicians and researchers in preventive medicine, statisticians, social scientists, and specialists in areas such as cardiovascular disease, cancer, infectious disease, etc.

IMPACT FACTOR

2018: 2.550 © Clarivate Analytics Journal Citation Reports 2019

ABSTRACTING AND INDEXING

PubMed/Medline BIOSIS Citation Index Embase CAB International Current Advances in Ecological and Environmental Sciences Current Contents - Clinical Medicine Research Alert Web of Science Cambridge Scientific Abstracts Scopus

EDITORIAL BOARD

AUTHOR INFORMATION PACK 17 Sep 2019

www.elsevier.com/locate/annepidem

GUIDE FOR AUTHORS

ACE Abstract Submission: Please see the full instructions here. Abstract submission is open until May 7, 2018. Late-breaker abstract submission closes on August 1, 2018.

Your Paper Your Way

We now differentiate between the requirements for new and revised submissions. You may choose to submit your manuscript as a single Word or PDF file to be used in the refereeing process. Only when your paper is at the revision stage, will you be requested to put your paper in to a 'correct format' for acceptance and provide the items required for the publication of your article.

To find out more, please visit the Preparation section below.

INTRODUCTION

Annals of Epidemiology is a peer reviewed, international journal devoted to epidemiologic research and methodological development. The journal emphasizes the application of epidemiologic methods to issues that affect the distribution and determinants of human illness in diverse contexts. Its primary focus is on chronic and acute conditions of diverse etiologies and of major importance to clinical medicine, public health, and health care delivery. *Annals* encourages the use of epidemiology in a multidisciplinary approach to understanding disease etiology. The *Annals* publishes Original articles, Brief Communications, Rapid Reports, Reviews, Commentaries, Editorials, Letters to the Editor, and Book Reviews.

Types of paper

The *Annals* follows the general instructions established by the "Uniform Requirements for Manuscripts Submitted to a Biomedical Journal". Specific guidelines for the different types of manuscripts are as follows:

Original article: a new piece of work-- <=3000 words; structured abstract <=200 words (Purpose, Methods, Results, Conclusions); <=6 Tables, Figures, or other illustrative materials.

Brief Communication: a short Original Article--<=2000 words; structured abstract <=200 words (Purpose, Methods, Results, Conclusions), <=3 Tables, Figures, or other illustrative materials. For Article Type at submission, please select Short Communication, then Brief Communication as the manuscript category.

Rapid Reports: a new fact, method, observation for quick presentation–<=1000 words; no abstract; maximum of one table, figure, or other illustrative materials. For Article Type at submission, please select Short Communication, then Rapid Reports as the

manuscript category.

Study Method and Design: a methodologic, logistic, or design focus that may or may not contain results but that presents something of interest in the design itself; simple descriptions of newly established studies, without results or an issue of intrinsic interest, are discouraged—format the same as that for original articles. For Article Type at submission, please select Original Article.

Reviews and Commentaries: a descriptive, analytic or synthetic consideration of a topic or hypothesis; word limit is negotiable, but in general should be <= 5000 words; unstructured abstract <=200 words; illustrative materials as needed.

Editorials: generally invited, but may be author initiated after consultation with Editor; usually ~1500 words directed to a paper in the same issue.

Letters to the Editor: a comment on a paper that has appeared recently in the Annals; generally ~750 words; no abstract.

Book Reviews: generally invited; usually ~1500 words or less.

Contact details for submission

Authors with special issues and concerns should contact the Editorial office. Authors are encouraged to include the name, address, phone/fax, and e-mail address of from one to three individuals they feel would provide an expert and careful review.

The *Annals* also publishes regular contributions from the *American College of Epidemiology*. Persons interested in contributing such materials—usually news and events of interest to epidemiologists—should contact the Editorial Office or the College directly.

All editorial questions and correspondence should be addressed to:
Annals of Epidemiology, C/O Patrick Sullivan & Cory Woodyatt Rollins School of Public Health, Emory University

U1518 Clifton Road NE, 4th Floor Atlanta. GA 30322 Editorial office's e-mail address: annals.of.epidemiology@emory.edu telephone: 404-692-2603

Fedex address: Urban Life Bldg Rm 848 140 Decatur St. Atlanta, GA 30303

Submission checklist

You can use this list to carry out a final check of your submission before you send it to the journal for review. Please check the relevant section in this Guide for Authors for more details.

Ensure that the following items are present:

One author has been designated as the corresponding author with contact details: • E-mail address • Full postal address

All necessary files have been uploaded: *Manuscript*: • Include keywords • All figures (include relevant captions) • All tables (including titles, description, footnotes) • Ensure all figure and table citations in the text match the files provided • Indicate clearly if color should be used for any figures in print *Graphical Abstracts / Highlights files* (where applicable)

Supplemental files (where applicable)

Further considerations • Manuscript has been 'spell checked' and 'grammar checked' • All references mentioned in the Reference List are cited in the text, and vice versa • Permission has been obtained for use of copyrighted material from other sources (including the Internet) • A competing interests statement is provided, even if the authors have no competing interests to declare • Journal policies detailed in this guide have been reviewed • Referee suggestions and contact details provided, based on journal requirements

For further information, visit our Support Center. **BEFORE YOU BEGIN**

Ethics in publishing

Please see our information pages on Ethics in publishing and Ethical guidelines for journal publication.

Studies in humans and animals

If the work involves the use of human subjects, the author should ensure that the work described has been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans. The manuscript should be in line with the Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals and aim for the inclusion of representative human populations (sex, age and ethnicity) as per those recommendations. The terms sex and gender should be used correctly.

Authors should include a statement in the manuscript that informed consent was obtained for experimentation with human subjects. The privacy rights of human subjects must always be observed.

All animal experiments should comply with the ARRIVE guidelines and should be carried out in accordance with the U.K. Animals (Scientific Procedures) Act, 1986 and associated guidelines, EU Directive 2010/63/EU for animal experiments, or the National Institutes of Health guide for the care and use of Laboratory animals (NIH Publications No. 8023, revised 1978) and the authors should clearly indicate in the manuscript that such guidelines have been followed. The sex of animals must be indicated, and where appropriate, the influence (or association) of sex on the results of the study.

Conflict of Interest

All authors must disclose any financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work. Examples of potential conflicts of interest include employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/ registrations, and grants or other funding. See also https://www.elsevier.com/conflictsofinterest. Download Conflict of Interest form (Conflict_of_Interest).

Submission declaration and verification

Submission of an article implies that the work described has not been published previously (except in the form of an abstract, a

published lecture or academic thesis, see 'Multiple, redundant or concurrent publication' for more information), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright- holder. To verify originality, your article may be checked by the originality detection service Crossref Similarity Check.

Preprints

Please note that preprints can be shared anywhere at any time, in line with Elsevier's sharing policy. Sharing your preprints e.g. on a preprint server will not count as prior publication (see 'Multiple, redundant or concurrent publication' for more information).

Use of inclusive language

Inclusive language acknowledges diversity, conveys respect to all people, is sensitive to differences, and promotes equal opportunities. Articles should make no assumptions about the beliefs or commitments of any reader, should contain nothing which might imply that one individual is superior to another on the grounds of race, sex, culture or any other characteristic, and should use inclusive language throughout. Authors should ensure that writing is free from bias, for instance by using 'he or she', 'his/her' instead of 'he' or 'his', and by making use of job titles that are free of stereotyping (e.g. 'chairperson' instead of 'chairman' and 'flight attendant' instead of 'stewardess').

Authorship

All authors should have made substantial contributions to all of the following: (1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, (3) final approval of the version to be submitted.

Changes to authorship

Authors are expected to consider carefully the list and order of authors **before** submitting their manuscript and provide the definitive list of authors at the time of the original submission. Any

addition, deletion or rearrangement of author names in the authorship list should be made only **before** the manuscript has been accepted and only if approved by the journal Editor. To request such a change, the Editor must receive the following from the **corresponding author**: (a) the reason for the change in author list and (b) written confirmation (e-mail, letter) from all authors that they agree with the addition, removal or rearrangement. In the case of addition or removal of authors, this includes confirmation from the author being added or removed.

Only in exceptional circumstances will the Editor consider the addition, deletion or rearrangement of authors **after** the manuscript has been accepted. While the Editor considers the request, publication of the manuscript will be suspended. If the manuscript has already been published in an online issue, any requests approved by the Editor will result in a corrigendum.

Copyright

Upon acceptance of an article, authors will be asked to complete a 'Journal Publishing Agreement' (see more information on this). An e-mail will be sent to the corresponding author confirming receipt of the manuscript together with a 'Journal Publishing Agreement' form or a link to the online version of this agreement.

Subscribers may reproduce tables of contents or prepare lists of articles including abstracts for internal circulation within their institutions. Permission of the Publisher is required for resale or distribution outside the institution and for all other derivative works, including compilations and translations. If excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article. Elsevier has preprinted forms for use by authors in these cases.

For gold open access articles: Upon acceptance of an article, authors will be asked to complete an 'Exclusive License Agreement' (more information). Permitted third party reuse of gold open access articles is determined by the author's choice of user license.

Author rights

As an author you (or your employer or institution) have certain rights to reuse your work. More information.

Elsevier supports responsible sharing

Find out how you can share your research published in Elsevier journals.

Role of the funding source

You are requested to identify who provided financial support for the conduct of the research and/or preparation of the article and to briefly describe the role of the sponsor(s), if any, in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. If the funding source(s) had no such involvement then this should be stated.

Funding body agreements and policies

Elsevier has established a number of agreements with funding bodies which allow authors to comply with their funder's open access policies. Some funding bodies will reimburse the author for the gold open access publication fee. Details of existing agreements are available online. After acceptance, open access papers will be published under a noncommercial license. For authors requiring a commercial CC BY license, you can apply after your manuscript is accepted for publication.

Open access

This journal offers authors a choice in publishing their research:

Subscription

• Articles are made available to subscribers as well as developing countries and patient groups through our universal access programs. • No open access publication fee payable by authors. • The Author is entitled to post the accepted manuscript in their institution's repository and make this public after an embargo period (known as green Open Access). The published journal article cannot be shared publicly, for example on ResearchGate or Academia.edu, to ensure the sustainability of peer- reviewed research in journal

publications. The embargo period for this journal can be found below. ***Gold open access***

• Articles are freely available to both subscribers and the wider public with permitted reuse. • A gold open access publication fee is payable by authors or on their behalf, e.g. by their research funder or institution.

Regardless of how you choose to publish your article, the journal will apply the same peer review criteria and acceptance standards.

For gold open access articles, permitted third party (re)use is defined by the following Creative Commons user licenses:

AUTHOR INFORMATION PACK 17 Sep 2019
www.elsevier.com/locate/annepidem 6

Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

For non-commercial purposes, lets others distribute and copy the article, and to include in a collective work (such as an anthology), as long as they credit the author(s) and provided they do not alter or modify the article.

The gold open access publication fee for this journal is **USD 3000**, excluding taxes. Learn more about Elsevier's pricing policy: https://www.elsevier.com/openaccesspricing.

Green open access

Authors can share their research in a variety of different ways and Elsevier has a number of green open access options available. We recommend authors see our open access page for further information. Authors can also self-archive their manuscripts immediately and enable public access from their institution's repository after an embargo period. This is the version that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and in editor-author communications. Embargo period: For subscription articles, an appropriate amount of time is needed for journals to deliver value to subscribing customers before an article becomes freely available to the public. This is the embargo period and it

begins from the date the article is formally published online in its final and fully citable form. Find out more.

This journal has an embargo period of 12 months.

Elsevier Researcher Academy

Researcher Academy is a free e-learning platform designed to support early and mid-career researchers throughout their research journey. The "Learn" environment at Researcher Academy offers several interactive modules, webinars, downloadable guides and resources to guide you through the process of writing for research and going through peer review. Feel free to use these free resources to improve your submission and navigate the publication process with ease.

Language (usage and editing services)

Please write your text in good English (American or British usage is accepted, but not a mixture of these). Authors who feel their English language manuscript may require editing to eliminate possible grammatical or spelling errors and to conform to correct scientific English may wish to use the English Language Editing service available from Elsevier's Author Services.

Informed consent and patient details

Studies on patients or volunteers require ethics committee approval and informed consent, which should be documented in the paper. Appropriate consents, permissions and releases must be obtained where an author wishes to include case details or other personal information or images of patients and any other individuals in an Elsevier publication. Written consents must be retained by the author but copies should not be provided to the journal. Only if specifically requested by the journal in exceptional circumstances (for example if a legal issue arises) the author must provide copies of the consents or evidence that such consents have been obtained. For more information, please review the Elsevier Policy on the Use of Images or Personal Information of Patients or other Individuals. Unless you have written permission from the patient (or, where applicable, the next of kin), the personal details of any patient included in any part of the article and in any supplementary materials (including all illustrations and videos) must be removed before submission.

Submission

Our online submission system guides you stepwise through the process of entering your article details and uploading your files. The system converts your article files to a single PDF file used in the peer-review process. Editable files (e.g., Word, LaTeX) are required to typeset your article for final publication. All correspondence, including notification of the Editor's decision and requests for revision, is sent by e-mail.

Submit your article

Please submit your article via https://www.evise.com/profile/api/navigate/AEP.

Referees

Please submit the names and institutional e-mail addresses of several potential referees. For more details, visit our Support site. Note that the editor retains the sole right to decide whether or not the suggested reviewers are used.

Manuscript Order

A cover letter should accompany the manuscript, addressing conflict of interest, authorship and nonduplication criteria. Please include suggestions of possible peer reviewers, as well. Submit entire manuscript on line in the following order: (1) title page; (2) abstract; (3) list of abbreviations and acronyms; (4) text; (5) acknowledgments and funding; (6) references; (7) legends; (8) tables; (9) figures. Number all pages in sequence, beginning with the title page as 1, abstract as 2, etc. In a cover letter indicate the single category the manuscript fits best.

Desk Rejections

When a manuscript is received, a initial triage is conducted by the Editor-in-Chief and Managing Editor to identify manuscripts that fit the journal's scope. In some cases, manuscripts may also be reviewed at this stage by an Associate Editor with subject matter

expertise in the area addressed by the manuscript. For details on why submissions are commonly rejected at this stage, please view our Desk Rejection Statement.

PREPARATION

NEW SUBMISSIONS

Submission to this journal proceeds totally online and you will be guided stepwise through the creation and uploading of your files. The system automatically converts your files to a single PDF file, which is used in the peer-review process. As part of the Your Paper Your Way service, you may choose to submit your manuscript as a single file to be used in the refereeing process. This can be a PDF file or a Word document, in any format or lay- out that can be used by referees to evaluate your manuscript. It should contain high enough quality figures for refereeing. If you prefer to do so, you may still provide all or some of the source files at the initial submission. Please note that individual figure files larger than 10 MB must be uploaded separately.

References

There are no strict requirements on reference formatting at submission. References can be in any style or format as long as the style is consistent. Where applicable, author(s) name(s), journal title/ book title, chapter title/article title, year of publication, volume number/book chapter and the article number or pagination must be present. Use of DOI is highly encouraged. The reference style used by the journal will be applied to the accepted article by Elsevier at the proof stage. Note that missing data will be highlighted at proof stage for the author to correct.

Formatting requirements

There are no strict formatting requirements but all manuscripts must contain the essential elements needed to convey your manuscript, for example Abstract, Keywords, Introduction, Materials and Methods, Results, Conclusions, Artwork and Tables with Captions. If your article includes any Videos and/or other Supplementary material, this should be included in your initial submission for peer review purposes.

Divide the article into clearly defined sections.

Figures and tables embedded in text

Please ensure the figures and the tables included in the single file are placed next to the relevant text in the manuscript, rather than at the bottom or the top of the file. The corresponding caption should be placed directly below the figure or table.

Peer review

This journal operates a single blind review process. All contributions will be initially assessed by the editor for suitability for the journal. Papers deemed suitable are then typically sent to a minimum of two independent expert reviewers to assess the scientific quality of the paper. The Editor is responsible for the final decision regarding acceptance or rejection of articles. The Editor's decision is final. More information on types of peer review.

REVISED SUBMISSIONS

Use of word processing software

Regardless of the file format of the original submission, at revision you must provide us with an editable file of the entire article. Keep the layout of the text as simple as possible. Most formatting codes will be removed and replaced on processing the article. The electronic text should be prepared in a way very similar to that of conventional manuscripts (see also the Guide to Publishing with Elsevier). See also the section on Electronic artwork.

To avoid unnecessary errors you are strongly advised to use the 'spell-check' and 'grammar-check' functions of your word processor.

Article structure

Subdivision - unnumbered sections

Divide your article into clearly defined sections. Each subsection is given a brief heading. Each heading should appear on its own separate line. Subsections should be used as much as possible when cross- referencing text: refer to the subsection by heading as opposed to simply 'the text'.

Introduction

State the objectives of the work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.

Material and methods

Provide sufficient details to allow the work to be reproduced by an independent researcher. Methods that are already published should be summarized, and indicated by a reference. If quoting directly from a previously published method, use quotation marks and also cite the source. Any modifications to existing methods should also be described.

Results

Results should be clear and concise.

Discussion

This should explore the significance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate. Avoid extensive citations and discussion of published literature.

Conclusions

The main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a Discussion or Results and Discussion section.

Appendices

If there is more than one appendix, they should be identified as A, B, etc. Formulae and equations in appendices should be given separate numbering: Eq. (A.1), Eq. (A.2), etc.; in a subsequent appendix, Eq. (B.1) and so on. Similarly for tables and figures: Table A.1; Fig. A.1, etc.

Essential title page information

• ***Title.*** Concise and informative. Titles are often used in information-retrieval systems. Avoid abbreviations and formulae where possible. • ***Author names and affiliations.*** Please clearly indicate the given name(s) and family name(s) of each author and

check that all names are accurately spelled. You can add your name between parentheses in your own script behind the English transliteration. Present the authors' affiliation addresses (where the actual work was done) below the names. Indicate all affiliations with a lower- case superscript letter immediately after the author's name and in front of the appropriate address. Provide the full postal address of each affiliation, including the country name and, if available, the e-mail address of each author.

• ***Corresponding author.*** Clearly indicate who will handle correspondence at all stages of refereeing and publication, also post-publication. This responsibility includes answering any future queries about Methodology and Materials. **Ensure that the e-mail address is given and that contact details are kept up to date by the corresponding author.**

• ***Present/permanent address.*** If an author has moved since the work described in the article was done, or was visiting at the time, a 'Present address' (or 'Permanent address') may be indicated as a footnote to that author's name. The address at which the author actually did the work must be retained as the main, affiliation address. Superscript Arabic numerals are used for such footnotes.

Highlights

Highlights are optional yet highly encouraged for this journal, as they increase the discoverability of your article via search engines. They consist of a short collection of bullet points that capture the novel results of your research as well as new methods that were used during the study (if any). Please have a look at the examples here: example Highlights.

Highlights should be submitted in a separate editable file in the online submission system. Please use 'Highlights' in the file name and include 3 to 5 bullet points (maximum 85 characters, including spaces, per bullet point).

Abstract

An abstract of 200 or fewer words should contain the following

clearly labeled elements: purpose, methods, results, and conclusions.

Keywords

Below the abstract 3 - 10 MeSH heading key words should be provided. Mesh headings may be found at http://www.nlm.nih.gov/mesh/meshhome.html . Editorials and Letters-to-the-Editor do not require abstracts.

List of abbreviations

Please list and define all abbreviations and acronyms used in the manuscript.

Acknowledgements

Collate acknowledgements in a separate section at the end of the article before the references and do not, therefore, include them on the title page, as a footnote to the title or otherwise. List here those individuals who provided help during the research (e.g., providing language help, writing assistance or proof reading the article, etc.).

Formatting of funding sources

List funding sources in this standard way to facilitate compliance to funder's requirements:

Funding: This work was supported by the National Institutes of Health [grant numbers xxxx, yyyy]; the Bill & Melinda Gates Foundation, Seattle, WA [grant number zzzz]; and the United States Institutes of Peace [grant number aaaa].

It is not necessary to include detailed descriptions on the program or type of grants and awards. When funding is from a block grant or other resources available to a university, college, or other research institution, submit the name of the institute or organization that provided the funding.

If no funding has been provided for the research, please include the following sentence:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Footnotes

Footnotes should be used sparingly. Number them consecutively throughout the article. Many word processors build footnotes into the text, and this feature may be used. Should this not be the case, indicate the position of footnotes in the text and present the footnotes themselves separately at the end of the article.

Artwork

Electronic artwork General points • Make sure you use uniform lettering and sizing of your original artwork. • Preferred fonts: Arial (or Helvetica), Times New Roman (or Times), Symbol, Courier. • Number the illustrations according to their sequence in the text. • Use a logical naming convention for your artwork files. • Indicate per figure if it is a single, 1.5 or 2-column fitting image. • For Word submissions only, you may still provide figures and their captions, and tables within a single file at the revision stage. • Please note that individual figure files larger than 10 MB must be provided in separate source files. A detailed guide on electronic artwork is available. **You are urged to visit this site; some excerpts from the detailed information are given here.** *Formats* Regardless of the application used, when your electronic artwork is finalized, please 'save as' or convert the images to one of the following formats (note the resolution requirements for line drawings, halftones, and line/halftone combinations given below): EPS (or PDF): Vector drawings. Embed the font or save the text as 'graphics'. TIFF (or JPG): Color or grayscale photographs (halftones): always use a minimum of 300 dpi. TIFF (or JPG): Bitmapped line drawings: use a minimum of 1000 dpi. TIFF (or JPG): Combinations bitmapped line/half-tone (color or grayscale): a minimum of 500 dpi is required. **Please do not:**

• Supply files that are optimized for screen use (e.g., GIF, BMP, PICT, WPG); the resolution is too low. • Supply files that are too low in resolution. • Submit graphics that are disproportionately large for the content.

Color artwork

Please make sure that artwork files are in an acceptable format (TIFF (or JPEG), EPS (or PDF), or MS Office files) and with the

correct resolution. If, together with your accepted article, you submit usable color figures then Elsevier will ensure, at no additional charge, that these figures will appear in color online (e.g., ScienceDirect and other sites) regardless of whether or not these illustrations are reproduced in color in the printed version. **For color reproduction in print, you will receive information regarding the costs from Elsevier after receipt of your accepted article**. Please indicate your preference for color: in print or online only. Further information on the preparation of electronic artwork.

Figure captions

Ensure that each illustration has a caption. A caption should comprise a brief title (**not** on the figure itself) and a description of the illustration. Keep text in the illustrations themselves to a minimum but explain all symbols and abbreviations used.

Tables

Please submit tables as editable text and not as images. Tables can be placed either next to the relevant text in the article, or on separate page(s) at the end. Number tables consecutively in accordance with their appearance in the text and place any table notes below the table body. Be sparing in the use of tables and ensure that the data presented in them do not duplicate results described elsewhere in the article. Please avoid using vertical rules and shading in table cells.

References

Citation in text

Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Any references cited in the abstract must be given in full. Unpublished results and personal communications are not recommended in the reference list, but may be mentioned in the text. If these references are included in the reference list they should follow the standard reference style of the journal and should include a substitution of the publication date with either 'Unpublished results' or 'Personal communication'. Citation of a reference as 'in press' implies that the item has been accepted for publication.

Reference links

Increased discoverability of research and high quality peer review are ensured by online links to the sources cited. In order to allow us to create links to abstracting and indexing services, such as Scopus, CrossRef and PubMed, please ensure that data provided in the references are correct. Please note that incorrect surnames, journal/book titles, publication year and pagination may prevent link creation. When copying references, please be careful as they may already contain errors. Use of the DOI is highly encouraged.

A DOI is guaranteed never to change, so you can use it as a permanent link to any electronic article. An example of a citation using DOI for an article not yet in an issue is: VanDecar J.C., Russo R.M., James D.E., Ambeh W.B., Franke M. (2003). Aseismic continuation of the Lesser Antilles slab beneath northeastern Venezuela. Journal of Geophysical Research, https://doi.org/10.1029/2001JB000884. Please note the format of such citations should be in the same style as all other references in the paper.

Web references

As a minimum, the full URL should be given and the date when the reference was last accessed. Any further information, if known (DOI, author names, dates, reference to a source publication, etc.), should also be given. Web references can be listed separately (e.g., after the reference list) under a different heading if desired, or can be included in the reference list.

Data references

This journal encourages you to cite underlying or relevant datasets in your manuscript by citing them in your text and including a data reference in your Reference List. Data references should include the following elements: author name(s), dataset title, data repository, version (where available), year, and global persistent identifier. Add [dataset] immediately before the reference so we can properly identify it as a data reference. The [dataset] identifier will not appear in your published article.

AUTHOR INFORMATION PACK 17 Sep 2019
www.elsevier.com/locate/annepidem 11

References in a special issue

Please ensure that the words 'this issue' are added to any references in the list (and any citations in the text) to other articles in the same Special Issue.

Reference management software

Most Elsevier journals have their reference template available in many of the most popular reference management software products. These include all products that support Citation Style Language styles, such as Mendeley. Using citation plug-ins from these products, authors only need to select the appropriate journal template when preparing their article, after which citations and bibliographies will be automatically formatted in the journal's style. If no template is yet available for this journal, please follow the format of the sample references and citations as shown in this Guide. If you use reference management software, please ensure that you remove all field codes before submitting the electronic manuscript. More information on how to remove field codes from different reference management software.

Users of Mendeley Desktop can easily install the reference style for this journal by clicking the following link: http://open.mendeley.com/use-citation-style/annals-of-epidemiology When preparing your manuscript, you will then be able to select this style using the Mendeley plug- ins for Microsoft Word or LibreOffice.

Reference formatting

There are no strict requirements on reference formatting at submission. References can be in any style or format as long as the style is consistent. Where applicable, author(s) name(s), journal title/ book title, chapter title/article title, year of publication, volume number/book chapter and the article number or pagination must be present. Use of DOI is highly encouraged. The reference style used by the journal will be applied to the accepted article by Elsevier at the proof stage. Note that missing data will be highlighted at proof stage for the author to correct. If you do wish to format the references yourself they should be arranged according to the following examples:

Reference Style

Follow the general arrangements as given in the “Uniform Requirements for Manuscripts Submitted to a Biomedical Journal”.

The journal uses the Vancouver style for references. Number references consecutively in the order in which they are first mentioned in the text. Cite references in text, tables and figures by arabic numerals in brackets (for example: text [2]. text [2, 3]). List all authors, but if the number exceeds six, list the first six followed by et al. Avoid using "abstracts," "unpublished observations" and "personal communications" as references. Papers accepted but not yet published may be included by stating the journal and adding "In Press." If such references are important to the understanding of the submitted manuscript, authors may include an electronic copy of the reference with their submission. The author is responsible for the accuracy of all references. References should be styled as follows:

Journals

[1] Paivio A, Jansen B, Becker LJ. Comparisons through the mind's eye. Cognition 1975;37(2):635-47. [2] Yuen AWC. Lamotrigine: a review of antiepileptic efficacy. Epilepsia 1994;35(Suppl. 5):S33-6. [3] VanDecar JC, Russo RM, James DE, Ambeh WB, Franke M. Aseismic continuation of the Lesser Antilles slab beneath northeastern Venezuela. J Geophys Res 2003;108:2043. https://doi.org/10.1029/2001JB000884.

Books

[4] Strunk Jr W, White EB. The elements of style. 3rd ed. New York: MacMillan; 1979 [chapter 4]. [5] College bound seniors. Princeton (NJ): College Board Publications; 1979. [6] Chaddock TE. Gastric emptying of a nutritionally balanced liquid diet. In: Daniel EE, editor. Proceedings of the fourth international symposium on gastrointestinal motility. Vancouver (British Columbia, Canada): Mitchell Press; 1974, p. 83-92.

Article or Chapter in an Edited Book

[7] Adams MJ, Briscoe BE, Sinha SK. Interface friction and energy dissipation in soft solid processing applications. In: Dowson D, Taylor CM, Childs THC, Godet M, Dalmas G, editors. Dissipative processes in tribology. Dowson D, editor. Tirbology series, vol. 27. Amsterdam: Elsevier; 1994, p. 223-34.

Edited Book

AUTHOR INFORMATION PACK 17 Sep 2019
www.elsevier.com/locate/annepidem 12

[8] Letheridge S, Cannon CR, editors. Bilingual education: teaching English as a second language. New York: Praeger; 1980.

Article in Press

[9] Chassin MR, Kosecoff J, Soloman DH. How coronary angiography is used. JAMA, in press.

Website

[10] Health Care Financing Administration. 1996 statistics at a glance,http://www.hcfa.gov/stats/stathili.htm; 1996 [accessed 13.03.12].

Dataset

[11] Oguro M, Imahiro S, Saito S, Nakashizuka T. Mortality data for Japanese oak wilt disease and surrounding forest compositions, Mendeley Data, v1; 2015.

Journal abbreviations source

Journal names should be abbreviated according to the List of Title Word Abbreviations.

Video

Elsevier accepts video material and animation sequences to support and enhance your scientific research. Authors who have video or animation files that they wish to submit with their article are strongly encouraged to include links to these within the body of the article. This can be done in the same way as a figure or table by referring to the video or animation content and noting in the body text where it should be placed. All submitted files should be properly labeled so that they directly relate to the video file's content. . In order to ensure that your video or animation material is directly usable, please provide the file in one of our recommended file formats with a preferred maximum size of 150 MB per file, 1 GB in total. Video and animation files supplied will be published online in the electronic version of your article in Elsevier Web products, including ScienceDirect. Please supply 'stills' with your files: you can choose any frame from the video or animation or make a separate image. These will be used instead of standard icons and will personalize the link to your video data. For more detailed instructions please visit our video instruction pages. Note: since

video and animation cannot be embedded in the print version of the journal, please provide text for both the electronic and the print version for the portions of the article that refer to this content.

Supplementary material

Supplementary material such as applications, images and sound clips, can be published with your article to enhance it. Submitted supplementary items are published exactly as they are received (Excel or PowerPoint files will appear as such online). Please submit your material together with the article and supply a concise, descriptive caption for each supplementary file. If you wish to make changes to supplementary material during any stage of the process, please make sure to provide an updated file. Do not annotate any corrections on a previous version. Please switch off the 'Track Changes' option in Microsoft Office files as these will appear in the published version.

Research data

This journal encourages and enables you to share data that supports your research publication where appropriate, and enables you to interlink the data with your published articles. Research data refers to the results of observations or experimentation that validate research findings. To facilitate reproducibility and data reuse, this journal also encourages you to share your software, code, models, algorithms, protocols, methods and other useful materials related to the project.

Below are a number of ways in which you can associate data with your article or make a statement about the availability of your data when submitting your manuscript. If you are sharing data in one of these ways, you are encouraged to cite the data in your manuscript and reference list. Please refer to the "References" section for more information about data citation. For more information on depositing, sharing and using research data and other relevant research materials, visit the research data page.

Data linking

If you have made your research data available in a data repository, you can link your article directly to the dataset. Elsevier collaborates with a number of repositories to link articles on ScienceDirect with relevant repositories, giving readers access to underlying data that gives them a better understanding of the research described.

There are different ways to link your datasets to your article. When available, you can directly link your dataset to your article by providing the relevant information in the submission system. For more information, visit the database linking page.

For supported data repositories a repository banner will automatically appear next to your published article on ScienceDirect.

In addition, you can link to relevant data or entities through identifiers within the text of your manuscript, using the following format: Database: xxxx (e.g., TAIR: AT1G01020; CCDC: 734053; PDB: 1XFN).

Mendeley Data

This journal supports Mendeley Data, enabling you to deposit any research data (including raw and processed data, video, code, software, algorithms, protocols, and methods) associated with your manuscript in a free-to-use, open access repository. During the submission process, after uploading your manuscript, you will have the opportunity to upload your relevant datasets directly to *Mendeley Data*. The datasets will be listed and directly accessible to readers next to your published article online.

For more information, visit the Mendeley Data for journals page.

Data statement

To foster transparency, we encourage you to state the availability of your data in your submission. This may be a requirement of your funding body or institution. If your data is unavailable to access or unsuitable to post, you will have the opportunity to indicate why during the submission process, for example by stating that the research data is confidential. The statement will appear with your published article on ScienceDirect. For more information, visit the Data Statement page.

AFTER ACCEPTANCE

Online proof correction

Corresponding authors will receive an e-mail with a link to our online proofing system, allowing annotation and correction of proofs online. The environment is similar to MS Word: in addition to editing text, you can also comment on figures/tables and answer questions from the Copy Editor. Web-based proofing provides a faster and less error-prone process by allowing you to directly type your corrections, eliminating the potential introduction of errors.

If preferred, you can still choose to annotate and upload your edits on the PDF version. All instructions for proofing will be given in the e-mail we send to authors, including alternative methods to the online version and PDF. We will do everything possible to get your article published quickly and accurately. Please use this proof only for checking the typesetting, editing, completeness and correctness of the text, tables and figures. Significant changes to the article as accepted for publication will only be considered at this stage with permission from the Editor. It is important to ensure that all corrections are sent back to us in one communication. Please check carefully before replying, as inclusion of any subsequent corrections cannot be guaranteed. Proofreading is solely your responsibility.

Offprints

The corresponding author will, at no cost, receive a customized Share Link providing 50 days free access to the final published version of the article on ScienceDirect. The Share Link can be used for sharing the article via any communication channel, including email and social media. For an extra charge, paper offprints can be ordered via the offprint order form which is sent once the article is accepted for publication. Both corresponding and co-authors may order offprints at any time via Elsevier's Author Services. Corresponding authors who have published their article gold open access do not receive a Share Link as their final published version of the article is available open access on ScienceDirect and can be shared through the article DOI link.

AUTHOR INQUIRIES

Visit the Elsevier Support Center to find the answers you need. Here you will find everything from Frequently Asked Questions to ways to get in touch. You can also check the status of your submitted article or find out when your accepted article will be published.